THE DEVELOPMENT OF GOVERNMENT CASH FORECASTING MODEL: A CASE STUDY FOR THE INDONESIAN GOVERNMENT

by
Iskandar Iskandar
BBsMn/BEcon, MSc (Econ)
Tasmanian School of Business and Economics

Submitted in fulfilment of the requirements for the degree of Doctor of Philosophy
University of Tasmania
June 2019

Declaration of Originality

This thesis contains no material which has been accepted for a degree or diploma by the University or any other institution, except by way of background information and duly acknowledged in the thesis, and to the best of my knowledge and belief no material previously published or written by another person except where due acknowledgement is made in the text of the thesis, nor does the thesis contain any material that infringes copyright.

Iskandar Iskandar
June, 2019

Authority of Access

- This thesis may be made available for loan and limited copying and communication in accordance with the Copyright Act 1968.
- This thesis may be made available for loan. Copying and communication of any part of this thesis is prohibited for two years from the date this statement was signed; after that time limited copying and communication is permitted in accordance with the Copyright Act 1968.
- This thesis is not to be made available for loan or copying for two years following the date this statement was signed. Following that time, the thesis may be made available for loan and limited copying and communication in accordance with the Copyright Act 1968.

Iskandar Iskandar
June, 2019

Acknowledgement

بِسۡمِ ٱللَّهِ ٱلرَّحۡمَٰنِ ٱلرَّحِيمِ

In the name of God (Allah), the Most Gracious, the Most Merciful. All praise is due to Allah, the Lord of the Worlds, on whom we depend for sustenance and guidance. I wish to acknowledge my sincere appreciation and gratitude to my supervisors, Professor Gary O'Donovan and Dr Shuxiang Xu for their invaluable guidance and patience throughout my PhD journey. I was lucky to have them both as my supervisors. I also would like to express my incomparable recognition to my lovely wife, Shinta Parlina, my marvellous children, Muhammad Khairullah faizin and Malika Kaisah Fathiyyah, my parents, and sisters for their pray, unconditional love, and unwavering patience. Without their support, finishing a PhD would be harder.

I wish to thank the Australia Awards Scholarships for providing me with the opportunity to do my research at the University of Tasmania. I also would like to acknowledge my high appreciation to the Directorate General of Treasury, Ministry of Finance, the Republic of Indonesia for the rewarding discussions and the welcoming access to their database such that vigorous information can be obtained by the present project.

Last but not least, special thanks to Professor Roger Willett for being my supervisor during the first two years of my study, my colleagues and staff at Tasmanian School of Business and Economics, University of Tasmania, my "Aussie" friends, brothers and sisters at Tasmanian University Union Muslim Society and Hobart Mosque, and all other persons who had manage to form a "support system" for me and my family during our fabulous time in Australia. It was a privilege to know you all.

Table of Contents

List of Figures

List of Tables

Abstract

The ability to predict the future cash required to fulfil government responsibilities and public services deliveries is crucial not only for the domestic economy but also for a potential spread to other communities. Discussions on the interconnection between government spending and economic development has been a prominent research area in the field of economic studies. However, the 2010 Greek crisis taught world a lesson that, regardless of existing causality, sustainable economic growth relies on the ways in which government manages expenditure. Moreover, public expenditure management (PEM) sees the national budget as an instrument to influence the economy through several features. One of them is cash management which focuses on ensuring the availability of government money to deliver public services in the most effective way. An effective government cash management (GCM) facilitates the requirements for the government to fulfil its responsibilities and public services deliveries while maintaining economic stability. Furthermore, a reliable government cash forecasting model is essential for an effective GCM.

In this thesis, the researcher has developed a government cash forecasting model that meets an acceptable level of accuracy and materiality for use by government cash managers. In doing so, the most appropriate variables were identified for use in the model and a number of statistical methods were evaluated and tested to be used to construct the model.

The methodology undertaken by this study was as follow. The government cash forecasting model developed utilised historical daily data of Indonesian government expenditure following three steps: (1) attribute selection, (2) modelling, and (3) performance evaluation processes. Several techniques based on statistical, machine learning, and hybrid methods were tested independently and then each was compared with the other to assist in developing the most accurate forecasting model based on performance evaluation measurements.

In the modelling phase, the following methods were used. The Autoregressive Integration Moving Average with Exogenous Variables (ARIMAX) technique was chosen to represent the statistical modelling method. The machine learning methods tested utilised multiple artificial neural network techniques including Feed-forward Neural Networks (FFNN), Cascade-forward Neural Networks (CFNN), Radial Basis Function Neural Network (RBFN), Generalised Regression Neural Network (GRNN), Extreme Learning Machine (ELM), Long Short-Term Memory (LSTM), and Gated Recurrent Unit (GRU). For the hybrid method, a combination of ARIMAX and Nonlinear Autoregressive Neural Network (NARNN) techniques was used.

The results show that an accurate government cash forecasting model that meets an acceptable level of materiality for the cash manager can be achieved by identifying and including the most significant variables which influence government expenditure through an attribute selection process and the application of an artificial neural network machine learning-based method. In this study, it was found that the most appropriate variables to build a government cash forecasting model are the total daily available fund for intermittent expenditure, the week of the month, the month of the year, and policy implementation, while the GRU was the most accurate technique.

This study contributes to the existing literature and practice in its development of a statistically robust and accurate method to forecast government expenditure. Notwithstanding that Indonesian data only was used in this research, the procedures used in this study and the forecasting model developed are applicable to other governments and public sectors.

Chapter 1 Introduction

The capacity to forecast the future of government cash requirements is essential not only to maintain the delivery of public services but also to achieve sustainable economic development. However, regardless of its importance, the government cash forecasting ability of many countries is limited. The 2010 Greek crisis is a case when the cost of borrowing became a burden to a government due to its failure to anticipate future government expenditure. Even where a government cash forecasting model is in place, if the model is not robust and accurate and does not match with the cash manager's expectations, then a failure to anticipate future government expenditure remains likely to occur. The present study is intended to explore the procedures that the cash manager may take to develop a government cash forecasting model with accuracy that meets an acceptable level of materiality for the cash manager.

The purpose of this chapter is to introduce the research presented in this study. The scope, context, and motivation of this study are conducted by summarising background information about the topic in Section 1.1. In Section 1.2., the rationale for the research is explained as a justification for the research which includes the purpose of the research and describing how the research aim (i.e. development of a government cash forecasting model) is supported by two research questions. A brief explanation of the methodology used to examine the research problem is summarised in the overview of the research as presented in Section 1.3. Lastly, the structure of the thesis is described in Section 1.4.

1.1. Background to the research

The cash management function has become the focal point of governments' activities due to their responsibilities in advancing and distributing public services (Widodo et al., 2014). Failure to manage government cash effectively might cause hazardous conditions to the domestic economy with potential spreads to wider global communities. The importance of advance government cash management is illustrated from the Greek crisis in 2010. Kouretas and Vlamis (2010) described the Greek emergency as a fiscal crisis, where the essential issues were brought about by unsustainable public debt. The revenues collected by the Greek government were not substantial to cover their spending. To compensate the shortfalls, the Greek government was forced to borrow more funds from the market (Kouretas & Vlamis, 2010). However, the strategy taken by the Greek government increased the cost of borrowing and debt burden since the investors were demanding higher interest rates (Arghyrou & Tsoukalas, 2011).

In general, there are two major economic theories explaining the linkage between government spending and economic development. The first theory is called Wagner's Law, referring to the work of Adolph Wagner. Wagner (1883) saw government spending as an endogenous variable to the economy (Tang, 2010). Following this law, the volume of public expenditure is driven by the size of its economy (Samudram et al., 2009). The second theory was proposed by John Maynard Keynes. In contrast to Wagner's Law, Keynes (1936) suggested government spending as an exogenous variable that affects its economic growth (Tang, 2010). Government expenditure is seen as the manifestation of fiscal policy in which the government might influence its economy. Therefore, the connection runs from government expenditure to economic growth (Samudram et al., 2009). However, regardless of the underlying theory underpinning the economy, the Greek crisis of 2010 shows that the way in which the government manages its expenditure is one of the keys for a steady economic development.

The interconnection between government expenditure and economic growth can also be seen from the point of view of Public Expenditure Management (PEM). According to Allen and Tommasi (2001), the focus discussion of PEM is in the area of the employment of economic resources in order to deliver government obligations and public services. From the perspective of PEM, the national budget is the main instrument for the government to manage the domestic economy (Allen & Tommasi, 2001). Therefore, the implementation of PEM can be seen from three aspects: budget preparation, budget execution, and cash management (Potter & Diamond, 1999). In the budget preparation process, macroeconomic indicators (e.g. inflation, exchange rates, economic growth) are set as a benchmark. The macroeconomic indicators are also seen as a target for the government to achieve during the budget execution, while cash management ensures that the provision of government money for public services delivery is done in the most effective way according to the macroeconomic framework (Allen & Tommasi, 2001).

Many studies of government cash management (GCM) have been undertaken with a focus on establishing an effective GCM (Mu, 2006; Lienert, 2009; Williams, 2010). Such studies agreed that the significant element of effective GCM is the government cash forecasting system. However, notwithstanding its essential role in effective GCM, accurate government cash forecasting in most countries is limited (Mu, 2006). Their cash forecasting models have failed to produce an accuracy that has satisfied the cash manager's expectations (Widodo et al., 2014; Comptroller and Auditor General, 2014). In order to strengthen their capacity for cash forecasting, Mu (2006) suggests that the cash manager analyses government cash expenditure patterns and builds a reliable cash forecasting model based on these patterns. In addition, most of the current literature is dominated by research in private corporate settings with little being concerned with the development of cash forecasting by the government.

In general, there are two techniques to build government cash forecasting models which are bottom-up or top-down approaches (Williams, 2010, 2009). The bottom-up method utilises information collected individually from all spending units to build a cash flow forecasting model, while the top-down technique depends on historical values of government cash expenditure of all spending units stored in a cash management database.

The GCM reform as initiated by the Indonesian government in 2003 was considered to be a success (Widodo et al., 2014). The government cash forecasting system that has been implemented by the Indonesian cash manager mainly follows the bottom-up approach. A review conducted by the cash manager criticised the poor quality of the forecasting model due to the onerous requirements for spending units to report and update their cash projections (Widodo et al., 2014). Despite some improvement with respect to simplifying the cash forecasting mechanism, Widodo et al. (2014) advocated the top-down forecasting approach, using historical patterns to enhance the accuracy of projections of expenditure flows. In spite of its success in establishing an effective GCM, providing a model with an accurate government cash forecasting is still a challenge for the Indonesian government. Therefore, the present study focusses its attention on the top-down approach to develop a reliable government cash forecasting model by utilising that of the Indonesian government.

Literature focussing on how to construct a government cash forecasting model is rare or inconclusive ((e.g. Sumando et al. (2018) and Iskandar et al. (2018)). This study makes a contribution to the body of knowledge by filling the gap in the literature by investigating the procedures involved in developing government cash forecasting model following the top-down approach and ultimately the development of a reliable government cash forecasting model. A better forecasting ability will help the government to manage its cash effectively. Moreover, the present study explores the historical data of government expenditure which allows the research to analyse the patterns of government cash. As suggested by Mu (2006), studying the patterns of government cash is essential to fortify the government cash forecasting ability. Therefore, insights gained from this study may be of assistance to understand the patterns of government cash expenditure. Furthermore, the research design presented in this study provides a framework for the cash manager to develop a government cash forecasting model. The procedures presented in this study are applicable to other governments and public sectors.

1.2. Research Objective and Research Questions

The objective of this study is to investigate the best means to develop a government cash forecasting model with accuracy that meets an acceptable level of materiality for the cash manager and it utilises the Indonesian government's historical data of government expenditure for this purpose.

In order to accomplish its objective, the present study explores and aims to identify the best variables and techniques to be used to construct a government cash forecasting model. Therefore, the main research aim for this study is:

To develop a government cash forecasting model with accuracy that meets an acceptable level of materiality for the cash manager by utilising historical government expenditure data.

The research aim can be addressed by answering questions about how to identify the best variables and techniques to be used to construct a government cash forecasting model. These are expressed as:

1. What are the most appropriate variables to be included in developing a government cash forecasting model with accuracy that meets an acceptable level of materiality for the cash manager by utilising the historical data of government expenditure?
2. What are the most appropriate techniques to use to develop a government cash forecasting model with accuracy that meets an acceptable level of materiality for the cash manager by utilising historical government expenditure data?

1.3. Research Methodology

In practice, there are two methods which are used in building a cash forecasting system, which are top-down analysis and bottom-up information (Williams, 2010, 2009). The GCM reform in Indonesia suggests that despite its successful implementation, providing a reliable government cash forecasting model is still problematic due to the poor quality of the forecasting model. Until recently, the government cash forecasting system in Indonesia has mainly followed the bottom-up approach with little attention to the top-down approach (Widodo et al., 2014). The case of the Indonesian government emphasises the importance of developing a government cash forecasting model following the top-down approach. Therefore, this study utilised the Indonesian government's expenditure historical data to investigate the best ways to develop a government cash forecasting model with accuracy that meets an acceptable level of materiality for the cash manager.

Several experiments were conducted in order to address the objective of the present study. Each experiment followed three sequential steps: (1) attribute selection stage; (2) modelling stage; and (3) performance evaluation stage. Firstly, all variables gathered from the data were verified based on their significance to influence the predictand, which was the government expenditure. At the conclusion of the attribute selection stage, two sets of data were chosen, namely the initial dataset and the selected attributes. The initial dataset consisted of all variables gathered from the data that influenced the predictand, while the selected attributes included only the significant one. Both data sets were utilised to construct a forecasting model in the modelling stage. The arrangement was intended to verify the effect of attribute selection on the performance of the model proposed by each data set. Therefore, the best variables to develop a government cash forecasting model was definitive.

In the modelling phase, multiple modelling techniques were employed to build a government cash forecasting model utilising both sets of data individually. The techniques represented various well-known methods in the area of forecasting research. The application of multiple modelling techniques as an independent model was deliberately prepared to identify the best technique to develop a government cash forecasting model with accuracy that meets an acceptable level of materiality for the cash manager.

Once the proposed datasets and techniques were modelled, the final stage was the performance evaluation phase. In this phase, all performances of the proposed models were measured based on different evaluation tools. These tools were carefully selected such that they represented most of the performance evaluation methods available in the theory. The purpose of the arrangement was to ensure the robustness of the selected model. The variables and technique employed by the selected model answered the research questions which ultimately address the study's research aim and questions.

1.4. The structure of the thesis

This thesis is structured as follows. A review of the literature relevant to the topic is presented in Chapter 2. Adopting the funnel-shaped approach, the literature is introduced from the broader area of PEM to a more specific scope of GCM and government cash forecasting. In the last section of Chapter 2, the various methods used by researchers in different forecasting problem settings are reviewed before the previous studies in the area of government cash forecasting model are discussed. The fact that the Indonesian cash manager is struggling to provide a reliable government cash forecasting model while successfully reforming its GCM, justifies the use of Indonesian government expenditure data as the case for this study. To provide a background understanding of the context of the data used in this study, the GCM and government cash forecasting practice in Indonesia is overviewed in Chapter 3. The research questions posed in the study are developed in Chapter 4. Addressing the research questions, multiple experiments are presented following the three sequential stages as discussed in Chapter 5. The experiments and the discussions of the results are presented in Chapter 6. Lastly, Chapter 7 concludes the present study with analysis of the results, contribution and limitations of the research and opportunities for future research.

Chapter 2 Literature Review

2.1. Introduction

The present study aims to develop a government cash forecasting model with accuracy that meets an acceptable level of materiality for the cash manager. As for the case of this study, a dedicated unit under the Ministry of Finance, namely the Directorate General of Treasury is acting as the Indonesian cash manager. In order to create a foundation for the research reported in this study, this chapter provides a review of literature relevant to the topic. The literature is introduced from the broader area of public expenditure management to a more specific scope of government cash management and government cash forecasting such that the logic of the study is easy to follow. The underlying framework for the present study is summarised in Figure 2.1.

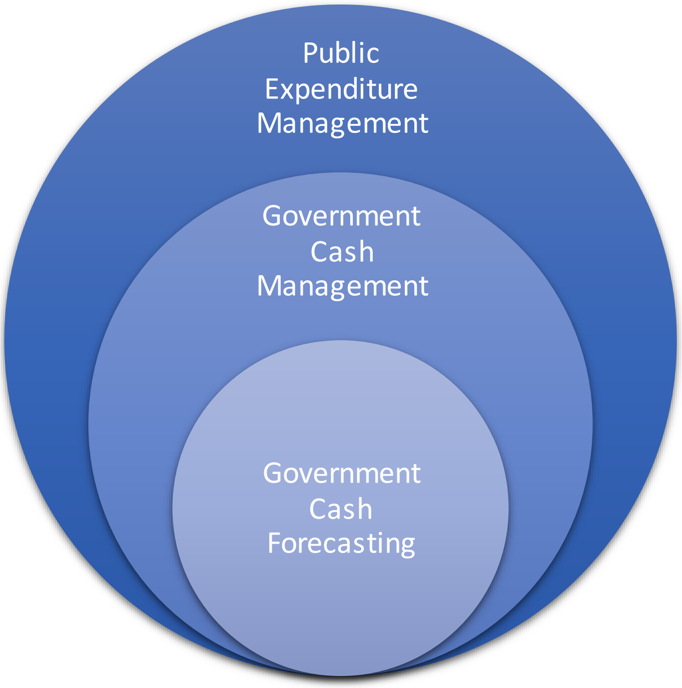


Figure 2.1: Framework for development of a cash forecasting model

Potter and Diamond (1999) considered the three aspects of public expenditure management that might affect the economy: budget preparation, budget execution, and cash management. Among those three, government cash management emerges to become a dominant function of all governments across the globe due to the role of government in promoting and delivering efficient public services (Widodo et al., 2014). The Greek crisis in 2010 is a good example where the failure of the Greek government in managing their cash effectively led to instability not only for the Greek economy but also with the possibility of a contagion effect spreading to other economies in the region (Arghyrou & Tsoukalas, 2011).

Moreover, Mu (2006) argued that an effective government cash management function can be achieved by establishing three functional building blocks: management of government cash receipts and payments, cash flow forecasting system, and management of government cash balance. The main purpose of the cash flow forecasting system is to provide the cash manager with an effective cash forecasting ability which is reliable and allows prediction of daily cash in-flows and out-flows of the government (Mu, 2006). Therefore, the need for a government cash forecasting model with accuracy that meets an acceptable level of materiality for the cash manager is an inevitable requirement for an effective government cash management function. In this study, the development of a government cash forecasting model focuses on predicting government payments made by all government spending units in Indonesia.

The purpose of this literature review chapter is to emphasise and articulate three issues: public expenditure management, government cash management and government cash forecasting. Section 2.2 outlines the broad picture of public expenditure management as a basis for the present study. The definition of government cash management and how to achieve effective government cash management is discussed in Section 2.3, while the theoretical and technical aspects of government cash forecasting are elaborated in Section 2.4. Lastly, previous studies on developing forecasting models and a summary are presented in Section 2.5. and Section 2.6. respectively.

2.2. Public Expenditure Management

The role of the national budget in economic performance has been a subject of research for many decades. Allen and Tommasi (2001) considered the national budget to be the primary instrument used by governments to manage the economy, while Potter and Diamond (1999) considered PEM as the economic perspective of the government expenditure. Prior research has analysed the relationship between government expenditure and economic performance in various settings. Current research such as Chipaumire et al. (2014), Abrishami et al. (2013), and Tang (2010) modelled the total of government expenditures, while Ogujiuba and Ehigiamusoe (2014), Magazzino (2012), and Oluwatobi and Ogunrinola (2011) disaggregated the data in an attempt to determine the effects of each type of government expenditure on economic performance. Some researchers (e.g. Alshahrani and Alsadiq (2014), Chipaumire et al. (2014), Chude and Chude (2013), and Menyah and Wolde-Rufael (2013)) focussed their attention on a specific country, whereas others undertook comparative research among both developed and developing economies (e.g. Kuckuck (2012), Lamartina and Zaghini (2011), and Magazzino (2011)).

There are two major alternative economic theories of public spending. One is based on the work of Adolph Wagner, the other is due to John Maynard Keynes. According to Wagner (1883), government spending is an endogenous variable in the macro-economy. In this view, public expenditure is determined by the growth of national income. Keynes (1936), in contrast, treated public expenditure as an exogenous variable that can affect economic development in the short-run (Tang, 2010). The Greek crisis of 2010 shows that, whichever theory is correct, sustainable economic growth relies on the way in which the government manages its expenditure.

Allen and Tommasi (2001) argued that in order to carry out its obligations, the government is entitled to various rights. One of them is budgeting, which allows the government to collect resources from the economy and allocate and use those resources responsibly, efficiently, and effectively (Allen & Tommasi, 2001). According to Allen and Tommasi (2001), the scope of Public Expenditure Management (PEM) is concerned with allocating and using economic resources in order to fulfil the government's responsibilities. Regardless of the absence of its formal definition in the literature (Ghiasi, 2014), the description of PEM can be found in its characteristics.

It is important to note that PEM is different from public sector budgeting. PEM focusses on the outcomes generated from public money while traditional budgeting focusses on public money as inputs to budget systems (Campos, 2001, Schick, 1998). Moreover, Campos (2001) compares the role of both instruments in delivering outputs. While public sector budgeting evaluates the quality of budget solely from its alignment to the arranged mechanisms (e.g. such as procurement procedures and payment processes) to produce an output, the PEM considers a good budget as the one that improves the desired outcomes. Therefore, the PEM suggests the improvement on regulations and procedures when the budget mechanisms fail to deliver the desired outcomes (Campos, 2001). In this regard, the PEM needs to be distinguished from the public expenditure policy (Allen & Tommasi, 2001).

In essence, public expenditure policy is questioning "what is to be done" while PEM is focussing on "how to implement the policy". An effort to delineate a strict boundary between policy and implementation will lead to impractical policies and ultimately bad implementation (Allen & Tommasi, 2001). Nevertheless, Allen and Tommasi (2001) argued that it is essential to distinguish PEM from the public expenditure policy due to the different goals both instruments are trying to achieve. The mechanisms, techniques, skills, and data required for a good PEM are different from the requirement to formulate a good public expenditure policy (Allen & Tommasi, 2001). However, to maintain its focus, this study limits its discussion to PEM only.

There are three objectives of PEM: (1) to maintain aggregate fiscal discipline; (2) to allocate resources in accord with government priorities; and (3) to promote the efficient delivery of services (Schick, 1998, Allen & Tommasi, 2001, Campos, 2001). Aggregate fiscal discipline assures public expenditures align with revenues in sustainable limits (Campos, 2001). It requires the capacity to set up expenditure estimation based on realistic revenue projection and fiscal targets. Such estimation is set as a benchmark during the budget formulation process (Schick, 1998, Allen & Tommasi, 2001). This can be done by setting up an expenditure ceiling on the aggregate level of individual spending units (Allen & Tommasi, 2001) and establishing a macroeconomic and fiscal framework as the starting point of budget preparation (Campos, 2001).

The primary objective of every budget system is to have control over the total expenditure (Allen & Tommasi, 2001). Campos (2001) supported the statement by equating the instillation of aggregate fiscal discipline with the "tragedy of the commons", where the agents involved consider the available resources are unlimited and their actions in utilising the resources are irrelevant to the total number of available resources. However, in a budget system, the opposite exists. The portion of the budget required by a spending unit is accumulated into a government expenditure. It is, in fact, funded by limited resources. In the absence of aggregate fiscal discipline, unrestrained expenditure will lead into untenable shortfalls that translate into an unstable macroeconomic environment which hinders the government to provide its obligations and maintain sustainable economic growth (Campos, 2001).

While maintaining aggregate fiscal discipline is about making the budget "cake" with regards to all the available resources (total revenues and funding) and constraints (macroeconomic and fiscal framework), allocating resources is related to the way of sharing the budget "cake" in accordance with government's strategic priorities (Campos, 2001). Due to limited resources, government expenditure should be efficiently allocated via a budget system that allows reallocation from lesser to higher priorities and to more effective programs relative to the government's objectives (Schick, 1998; Allen & Tommasi, 2001). Although determining the best configuration between government priorities and effective programme is challenging, Campos (2001) suggested the involvement of the citizen in prioritising the government programmes and the line agency to estimate the cost of the programme is worthwhile. However, the available mechanism fails to reveal the citizen's true reference on the government programmes. Moreover, the spending units are reluctant to fully disclose information regarding the cost of the programme and tend to "play safe" by submitting the low-cost estimation for the programme in order to secure a spot in the budget system (Campos, 2001).

Promoting an efficient delivery of public services is focussing on the implementation of value for money principles in the public sector (Campos, 2001). It is a technical efficiency with concerns on the operational level to implement programmes within spending units at reasonable quality and the lowest cost possible on the basis of efficiency and effective management system (Allen & Tommasi, 2001; Campos, 2001). This includes two elements: (1) producing goods and delivering services at an efficient cost and (2) providing money to be used for producing and delivering public services (Potter & Diamond, 1999).

According to Allen and Tommasi (2001), these three objectives are complementary and interdependent. Strong fiscal discipline is a necessary condition to have an effective budget allocation with regard to government priorities. Enhancing the internal management system to achieve efficiency without strict budget constraints is meaningless. On the other hand, relying only on fiscal discipline while resource allocation and technical budget operations are inefficient, will lead to an unsustainable budget system. Establishing a budget ceiling without an effective public expenditure system will result in underfunding priorities and possibly conflict with the ability to implement policy. Moreover, Allen and Tommasi (2001) argued that the objectives of PEM can be seen as a framework to assess the progress of budget system reforms. To be successfully implemented, all three objectives of the PEM have to be achieved concurrently such that the progress toward one objective is not sacrificing the achievement of other objectives (Allen & Tommasi, 2001).

According to Potter and Diamond (1999), there are three critical aspects of public expenditure management: budget preparation, budget execution, and cash planning and management. While budget preparation mostly deals with macroeconomic indicators (e.g. inflation, exchange rates, economic growth) budget execution is mainly about expenditure procedures (e.g. commitment and procurement processes). Cash planning and management focus on ensuring the availability of government money to deliver public services in the most effective way (Allen & Tommasi, 2001).

Furthermore, Potter and Diamond (1999) elaborated the three objectives of PEM as previously described into four fiscal disciplines: (1) control of aggregate expenditure to ensure affordability, that is, consistency with the macroeconomic constraints; (2) effective means for achieving a resource allocation that reflects expenditure policy priorities; (3) efficient delivery of public services (productive efficiency); and (4) minimisation of the financial costs of budgetary management (i.e., efficient budget execution and cash and debt management practices). The interconnection between objectives, fiscal disciplines, and aspects of PEM is shown in Table 2.1.

Table 2.1: The interconnection between objectives, features, and phases of PEM

Objectives	Fiscal disciplines	Aspects
to maintain aggregate fiscal discipline	control of aggregate expenditure to ensure affordability, that is, consistency with the macroeconomic constraints;	budget preparation
to allocate resources in accord with government priorities	effective means for achieving a resource allocation that reflects expenditure policy priorities;	budget preparation
to promote the efficient delivery of services	efficient delivery of public services (productive efficiency); and	budget execution
	minimisation of the financial costs of budgetary management (i.e., efficient budget execution and cash and debt management practices).	cash planning and management

Source: Adapted from Campos (2001), Allen and Tommasi (2001), Potter and Diamond (1999)

The government decides the amount of public expenditure and budget allocation prior to the beginning of the budget year. This decision is taken with the consideration of macroeconomic targets and fiscal frameworks in terms of an expenditure ceiling and government programme prioritisation. Such consideration is necessary to ensure that the level of deficit which might exist can be safely financed without causing shock to the economy. In addition, macroeconomic projections help the government to determine their strategic priorities so that resources can be allocated efficiently (Potter & Diamond, 1999).

In the budget execution phase, spending units undertake activities such as commitment making, procurement, verification, and payment processes as part of their planned expenditure (Potter & Diamond, 1999). These activities are bound by a set of regulations and procedures enacted by the cash manager to make sure the budget is the best value for money. During this phase, in-year adjustments are permissible to accommodate policy changes and to maintain the quality of the budget and promote the efficient delivery of public services.

Another way of achieving efficient public services delivery is by minimising the cost of borrowing to cover in-year shortfalls of planned expenditure, which is one of the features of a government's cash management function. The cash manager needs to develop the skill to reduce expensive borrowing caused by unplanned cash expenditure. Potter and Diamond (1999) argued that government cash management is an indispensable phase in achieving successful PEM. It is also essential for a government to have an effective cash management plan to ensure unnecessary interruption to public services delivery is avoided (Potter & Diamond, 1999). The aim of this research and thesis is to develop a government cash forecasting model as a measure to achieve effective government cash management. An in-depth discussion on government cash management is presented in the following section.

2.3. Government Cash Management

Many studies have proposed a definition of government cash management (GCM). Storkey (2003) defined GCM as a condition such that the government's responsibilities can be effectively funded in the correct amount, in the exact location, and at the right time. Furthermore, Mu (2006) and Williams (2009) described GCM as a set of action plans and related procedures regarding the government's short-term cash flow and cash balances, between government institutions and between the government and other sectors. However, all definitions focus on the time value of government money which refers to minimising the cost of the funding by matching the time difference between when the money is needed and its availability (Lienert, 2009).

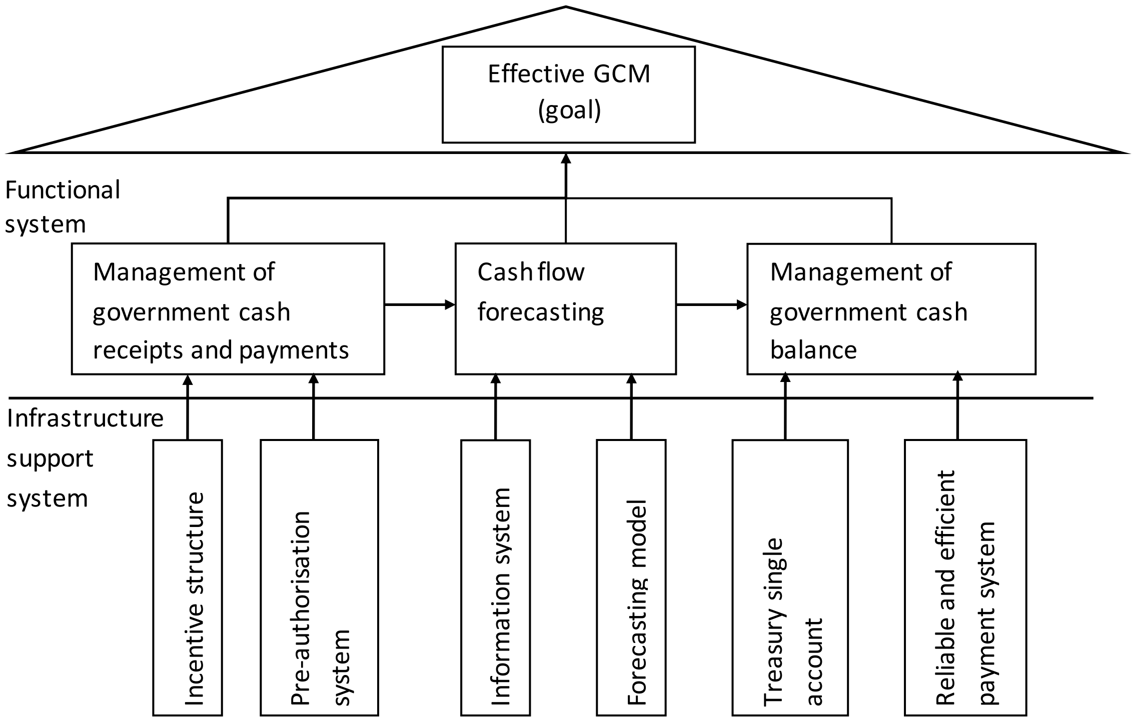
A clear understanding of GCM is essential for cash managers in the public sector. The cash managers need to distinguish GCM from government budget management (Mu, 2006, Williams, 2009). Most of the time the cash managers use a cash limitation policy in dealing with cash mismatches. It can be done by postponing the programme until the fund is available or cancelling it for the rest of the budget year. The policy can be implemented on spending units, economic categories of spending or even individual line items basis (Potter & Diamond, 1999). According to Williams (2009), reducing budget appropriations due to a lack of government money is 'cash rationing' which is part of budget management, not cash management.

The main objective of budget management is to make sure that the budget is executed within budget appropriations. It is part of budget management's authority to make an in-year adjustment on budget appropriations, if required. On the other hand, it is a duty for the cash manager to ensure that the government has the liquidity to complete its payments once budget appropriations are set. Instead of reducing the planned expenditure when a cash mismatch occurs, the cash manager may choose to borrow money. (Lienert, 2009). If cash managers fail to provide sufficient funds when they are needed, it leads to costly borrowing (Williams, 2010). This problem is alleviated by effective GCM.

Schiavo-Campo (2017) argued that the GCM function comprises (1) providing spending agencies with the funds needed to implement their activities in a timely fashion; (2) managing government cash to minimise the cost of government borrowing; (3) managing public debt; (4) managing the financial assets and liabilities of the government; and (5) handling all government bank accounts required for GCM purposes. Regardless of the budgeting system adopted by the government, an effective GCM is an essential feature to control macroeconomic and fiscal policy implementations (Schiavo-Campo, 2017). However, most developing countries put their focus onto the compliance of budget execution in accordance with its procedures and guidelines, which is part of budget management (Campos, 2001), rather than the cash management (Schiavo-Campo, 2017). In addition, the results of a survey by Schiavo-Campo (2017) indicated that substantial savings can be made by improving the management of government cash to minimise borrowing costs and maximise short-term returns.

A definition of effective GCM can be derived from its objective which is to have a certain amount of money available in a stated place at the time when it is needed to meet government liabilities at minimum cost (Mu, 2006). This is achieved through reliable forecasting of cash flows and balances, accountable cash management and service provision procedures at minimum risk, and the integration of a cash and debt management function (Storkey, 2003). Pooling all of the government's money in a Treasury Single Account (TSA) and establishing a cash flow forecasting model facilitates reliable future cash flows and balance projections. Modernising cash management's infrastructure (e.g. transaction processes, accounting frameworks, institutional arrangements and responsibilities, and information sharing) fosters the cash manager's responsibilities in every level and supports cost and risk reduction (Lienert, 2009). Close coordination between cash and debt management promotes more active cash management and minimises cost (Williams, 2010).

To ensure an effective GCM can be attained, Mu (2006) suggested three functional building blocks to be established. Those three functional building blocks – as shown in Figure 2.2. – are the Government Cash Receipt and Payment Management Function, the Government Cash Flow Forecasting Function, and the Government Cash Balance Management Function. Each function is sustained with its respective infrastructure support system. The cash manager employs two infrastructure systems – information system and forecasting model – to support the cash flow forecasting function.



Source: Adapted from Mu, (2006)

Figure 2.2: Building Blocks of an Effective Government Cash Management System

The cash manager utilises the information system to create a database of all historical data of cash inflows and outflows. The information gathered from the database is then input into a forecasting model in order to predict future cash needed. Furthermore, this database allows the cash manager to identify and understand cash patterns, useful to improve the performance of the forecasting model. Having an appropriate information system to collect government cash flow data and a reliable forecasting model provides the fundamental infrastructure for effective GCM (Mu, 2006).

There are four objectives of modern cash management: (1) to ensure government expenditures are made in a timely manner; (2) to ensure effective borrowing with minimum cost; (3) to maximise returns on idle cash; and (4) to manage risks (Lienert, 2009). Moreover, Lienert (2009) investigated key features of effective modern GCM practised in advanced OECD countries and identified the six fundamental features and three desirable features shown in Figure 2.3.

Fundamental features

1. Centralization of government cash balances and establishment of a Treasury Single Account (TSA) structure
2. Clear understandings on the coverage of the cash planning framework
3. Ability of make accurate projections of short-term cash inflows and outflows
4. An adequate transaction processing and accounting framework
5. Timely information sharing between the central treasury, revenue-collecting agencies, spending ministries, and/or treasury branch offices
6. Appropriate institutional arrangements and responsibilities

Desirable features

7. Utilization of modern banking, payment, and settlement systems
8. Use of short-term financial market instruments for cash management
9. Integration of debt and cash management

Source: Lienert (2009)

Figure 2.3: Nine Features of Modern Cash Management

Similarly to Lienert (2009), Williams (2010) elaborated Lienert's nine key features into five key characteristics of good practice in GCM, as listed in Figure 2.4.

- Centralization of government cash balances and establishment of a TSA
- Modern systems: an adequate transaction processing and accounting framework (processing government transactions with few handling steps, reliance on electronic transactions); modern banking, payment, and settlement systems
- Ability to make accurate projections of short-term cash inflows and outflows
- Strong institutional interaction, covering in particular:
  - Information sharing between the cash managers, revenue-collecting agencies and spending ministries (and any relevant ministry branch offices)
  - Strong coordination of debt and cash management
  - Formal agreements between the Ministry of Finance and the central bank on information flows and respective responsibilities
- Use of short-term instruments (treasury bills, repo and reverse repo, term deposits, etc.) to help manage balances and timing mismatches

Source: Williams (2010)

Figure 2.4: Key Characteristics of Good Practice in GCM

Lienert (2009) illustrated four steps for countries to follow in moving to more modern GCM. The first was establishing the basic principles for cash management reform. This included establishing a dedicated unit responsible for both operational and policy-making of GCM; underlining the importance of effective cash management on minimising unnecessary cost; constructing a credible annual budget structure; establishing a TSA to pool all revenues and to pay all expenditures from one operational bank account at the central bank; utilising the banking system to support cashless transactions for both revenues and expenditures; endorsing a more conservative cash advances mechanism; enhancing the quality of government accounting in which timely, comprehensive and high-frequency data is available for short-term cash predictions and amending the legal framework as needed to justify government banking arrangements; and the establishment of a TSA for instance (Lienert, 2009).

Apart from institutional arrangements amongst related agencies, the crucial point in the first step is the establishment of the TSA itself. Minimising the amount of government idle cash can be done by pooling all government cash balances into a TSA. Such an arrangement allows the cash manager to control previously undetected idle cash in one area of government activity. The TSA prevents the cash manager from acquiring unexpected costly borrowing to finance the cash shortages in another government spending unit. It does this by allowing the cash manager to use cash surpluses in one spending unit to cover the costs of government programmes in another (Williams, 2010). In addition, the TSA also provides historical data for cash forecasting purposes (Mu, 2006, Lienert, 2009). Despite Pattanayak and Fainboim (2010), suggesting that the main account of a TSA system may be held at a commercial bank, Williams (2010) suggested the central bank is the preferable place.

The arrangements in which the TSA interacts with the government payment process come in many forms (Williams, 2010). A centralised arrangement allows the cash manager to fully control cash outflows from a TSA by processing all the expenditure transactions made by spending units centrally (Pattanayak & Fainboim, 2010). On the other hand, the cash managers might give authorisation to spending units for expenditure payments from separate accounts in the banking system. In this case, the responsibility for processing government payments is on the central bank via interbank clearance systems or the banking sector under specific contracts (Williams, 2010). Some governments combine both arrangements where major payments are executed directly through the TSA and smaller transactions are paid via the commercial banking system. Regardless of the approach chosen by the cash manager, the key element of a TSA is that any balances left with the banking system overnight are transferred back into the TSA (Pattanayak & Fainboim, 2010, Williams, 2010).

The second step is preparing cash plans and developing cash management skills (Lienert, 2009). It is a duty of a cash manager to make preparations for cash planning by initiating short-term projections of cash inflows and outflows, evaluating its current performance, updating the database, and revising predictions in addition to preparing short-term cash flow projections, forming information-sharing arrangements and ensuring that information exchanges for cash projections take place. Such arrangements require a technical network between the cash manager, the revenue-collection agencies, major spending units, and the central bank to provide inputs for projections of short-term cash flows. Therefore, the cash manager can maintain first-hand information about regularly occurring and scheduled major cash flows which are relatively easy to predict. It is preferable to prepare cash plans progressively monthly, bi-monthly, or weekly with daily cash plans as the ultimate goal (Lienert, 2009). Moreover, Lienert (2009) emphasised the need for ongoing training for GCM operators such as projecting, monitoring, analysing, and updating cash plans as skilled staff are required to ensure cash management is optimised.

Information-sharing arrangements can be enhanced through the use of information technology (IT). High-performing IT systems such as integrated financial management information systems (IFMIS) might be needed to facilitate the preparation and updating of short-term cash projections and the maintaining of databases which report cash-flow trends (Lienert, 2009). IFMIS supplies the cash manager with comprehensive financial information, not only for the current budget year but also past periods, which is available for in-year budget controlling purposes and to improve economic forecasting, planning, and budgeting (Khemani & Diamond, 2005). However, for cash flow forecasting purposes, Williams (2010) argued that the establishment of IFMIS is not urgent due to the nature of the data needed. Cash flow predictions are an operational decision which does not require the data to be as precise as for accounting reporting since it is required for immediate operational policymaking. The cash manager might develop a separate database for short-term cash flow projections from IFMIS. The key points are that it has to be flexible and fully controlled by the cash manager (Williams, 2010).

Improving the basic requirements and cash planning was the next phase of the GCM reform proposed by Lienert (2009). This improvement encompassed a set of internal arrangements and the establishment of agreements with external parties. It is necessary to maintain solid coordination between the government revenue collection agencies and the cash manager to minimise time delays in accruing revenues into the TSA. On the other hand, synchronising cash and debt management policy is required in order to guarantee the bond issuance program in line with the cash profile, to hinder avoidable debt issuances, and to make sure the interconnection between the money and bond markets is consistent with the macroeconomy

From the technical side, improving expenditure approvals and payment procedures with the help of information and communication technology (ICT), broadening the TSA coverage to all government bank accounts including extra budgetary account, and minimising idle cash, are amongst the main foci at this stage. Reviewing revenue and expenditure regulations from the cash management's point of view to handle shortage in cash inflows and evaluating the implication of cash predictions to the planned expenditure commitments by reviewing government procurement procedures – when required – are necessary to do. Moreover, at this stage Lienert (2009) encouraged the cash manager to settle agreements with stakeholders outside the government such as the banking sector and the central bank. An arrangement to enhance the use of banking facilities to eliminate the time lag of government payments and sustain the cash balances in government bank accounts outside the TSA's main account at the central bank is necessary for more effective GCM. At the initial stage, the selection of commercial banks for the provision of transactional banking services to the treasury purposes needs to be stated in a formal contracting document based on a competitive bidding basis. Negotiating remuneration of government idle cash balances held by the commercial banks and the central bank is the next arrangement to make. Clarifying and improving the equal relationship and the consultation between the central bank as a monetary policymaker and the government cash manager covered in a memorandum of understanding (MoU) or a service level agreement (SLA) is advisable (Lienert, 2009).

In this step, the focus of the cash manager is to strengthen the implementation of government cash management reforms. A review of government expenditure procedures is necessary to ensure the fulfilment of government obligations and the delivery of public services are accomplished at minimum cost. Defining the national budget into the revenues and expenditures listed in the document of mutual agreement between government and legislature limits the disclosure of all government accounts (Cotterell & Wickens, 2007). All government accounts that are excluded from the budget system are classified as the extrabudgetary account by the International Monetary Fund (IMF) via the Government Finance Statistics (GFS) frameworks (IMF, 2014). In most countries, it is a common practice to have extrabudgetary accounts for special purposes such as social security funds, development funds, and reserve (Allen & Radev, 2010). Extending the TSA to cover the extrabudgetary accounts widens the opportunity for the cash manager to use the balance for cash management purposes such that unnecessary borrowing can be hindered (Lienert, 2009). It is not compulsory for an effective government cash management to have integrated cash and debt management functions. However, strong coordination between both functions is a necessary condition (Williams, 2010). Cash management and debt management units are independent bodies. A joint committee is desirable to bring such units together. In addition, an intensive daily interaction at policy and operational levels is crucial to ensure the operation of debt management unit is in line with the cash manager's policy such that the borrowing made to cover the cash shortages does not cause a burden for the government (Williams, 2010). Apart from the technical operational arrangements between the cash manager and the central bank, the agreement might include the remuneration of the idle government cash balances held by the central bank based on the market interest rates (Lienert, 2009).

Once GCM is developed, it is desirable for the cash manager to shift focus towards more active daily management as the last step (Lienert, 2009). According to Lienert (2009), more active daily cash management can be achieved via several arrangements. These arrangements are investing daily cash surpluses of a TSA in financial markets regardless of the existence of a daily operating target, arranging a daily sweeping of all government accounts into a TSA, protecting short-term placements by the cash manager from default risk, enhancing the accuracy of cash flow forecasting and reinforcing coordination between the cash manager, the government debt manager, and the central bank (Lienert, 2009).

There are many short-term instruments that can be used by the cash manager to manage the daily cash available such as treasury bills (T-Bills), repurchase agreement (repo), and reverse repo (Lienert, 2009). T-Bills are discounted money market instruments with maturity of less than a year (Williams, 2010). Similar to T-Bills, repo is also considered as a short-term borrowing instrument. In the repo, the cash manager sells securities with an agreement to buy them back later. On the other hand, under the reverse repo, the cash manager invests their temporary cash surplus to short-term securities with an agreement to sell them back on their maturity date (Lienert, 2009). Such instruments allow the cash manager to handle short-term temporal cash shortfalls and surpluses. When the forecasted expenditures are exceeding the predicted revenues, the cash manager issues T-Bills and repo. However, in the case where the cash surpluses exist, the reverse repo option can be chosen to optimise idle cash. In addition, repo and reverse repo are fully collateralised, thus, the cash manager can minimise the risk of borrowing or investment. While treasury bills are the basic short-term instruments, repo and reverse repo are considered as instruments for more active daily cash management due to policy imperatives (Williams, 2009). Essentially, treasury bills are the cash manager's tools to maintain government cash balance in the banking system such that the volume of the treasury bills issued is subject to inflows and outflows at a particular time of the budget year. The maturity of the treasury bills most likely spans a period of a couple of weeks up to several months. The treasury bills issuance is expected to be higher when the shortfall between inflows and outflows is predicted to be higher in the following week, for instance (Williams, 2010). On the other hand, a more active daily cash management policy intervenes in the monetary policy held by the central bank. As policy makers, both the cash manager and the central bank have different priorities and hazard sensitivities. In an absence of sufficient procedures and planning arrangements between both institutions, the risk of conflicting policy that leads to wider economic damage exists (Pessoa & Williams, 2012). Furthermore, the cash manager, together with the central bank, also interacts with the banking system as users of banking services or as a regulator. Therefore, when using instruments like repo and reverse repo, a number of agreements with the central bank and the banking sector need to be set up by the cash manager for market compliance (Williams, 2010).

In agreeing with Lienert (2009), Williams (2009) summarised the four stages into (1) the establishment of TSA; (2) the development of a cash forecasting system; (3) “rough tuning”; and (4) “fine tuning”. The government accounts consolidation and the sweeping of overnight balances into a single account held by the Treasury are the centre of the formation of a TSA. The cash manager’s ability to forecast and monitor the availability of government money is important to be able to make decisions that minimise costs from unnecessary borrowing. “Rough tuning” aims to ensure the government cash balances are manageable within the cash target set by the cash manager using the short-term money market instruments, most likely the T-bills, repo, and reverse repo. Lastly, the “fine tuning” ensures that more active GCM operations are not conflicting with other financial policy leading to shock in the economy. More detailed agreements with the central bank and the banking sector are needed for “fine tuning” the GCM (Williams, 2009).

Lienert (2009) and Williams (2010) agree on the urgency of improving cash forecasting to achieve effective GCM. Lienert (2009) argued that the necessary condition for effective cash management is that the cash manager has adequate ability to collect, maintain, and predict short-term cash inflows and outflows into government accounts. Cash flow forecasting is fundamental for more active cash management (Williams, 2010). Moreover, Williams (2010) suggested three months as a desirable period for which daily cash inflows and outflows should be able to be reliably forecasted.

This study investigates the procedures that cash managers might take in order to prepare better cash planning by initiating short-term predictions of cash outflows. Notwithstanding the fact that most research into government cash forecasting is complementary to research into government cash management, some studies elaborate on the key features required to establish a reliable government cash forecasting system. These studies are discussed in the following section.

2.4. Government Cash Forecasting

For effective cash management, the cash manager needs to accurately develop timely short-term estimates of cash inflows and outflows. The flows to be forecast include government receipts and payments (i.e., those that contribute to the fiscal balance—deficit or surplus) and financing transactions (i.e., changes in net financial assets and liabilities, which finance the fiscal balance). A key objective is to anticipate the cash needs of the government and to ensure that payments are made in a timely manner (Lienert, 2009).

Although government cash flow forecasting ability is critical for the cash manager, it is extremely weak in most developing countries (Mu, 2006). Most government spending units neglect the value of money, which potentially leads to sub-optimum decisions (Storkey, 2003). They are not aware that unremunerated idle cash might lead to an opportunity cost for the cash manager. However, a report released by the Comptroller and Auditor General (2014) indicated the challenge is also faced by developed countries where failure in providing accurate expenditure forecasting has led to substantial unprecedented costs. Moreover, excellent cash forecasting is a necessary condition in achieving value for money (Comptroller and Auditor General, 2014). Information regarding the timing and amount of government expenditure is required for the cash manager to be able to minimise opportunity cost. It can be provided with a reliable government cash forecasting system.

As a technical exercise, developing government cash forecasting demands active participation from all stakeholders including the cash manager, revenue-collecting agencies, and spending units (Lienert, 2009). Moreover, Lienert (2009) describes the preparatory measures that a cash manager should take, before establishing short-term cash flow forecasts. These include ensuring the daily, weekly and monthly cash flow forecasting is coherent with annual budget projections, focusses on the accuracy of major cash inflows and outflows, analyses historical patterns of certain cash flows, recognises when major flows occur, establishes information networks between the cash manager, revenue-collecting agencies, and spending units to ensure real-time cash flows projection updates, confirming spending units' expenditure commitments to enhance the accuracy of predicted cash needed and communicates effectively to spending units that the information provided to the cash manager is for cash management purposes, not for expenditure control purposes (Lienert, 2009). These prerequisite conditions are necessary to collect preliminary information required for government cash forecasting modelling development. Active participation from all stakeholders ensures the collected information is credible.

To strengthen the ability to accurately forecast cash flows, Mu (2006) suggests developing countries establish a database to collect historical data of inflows and outflows, analyse patterns of cash flows, and develop cash forecasting models from historical data. Mu (2006) also emphasises the necessity of observing time constrained events and information that affects cash flows (e.g. public holidays, payment dates on weekends, and payment lead times) and to enable making regular adjustments to forecasting models.

In practice, techniques used to design a cash forecasting system vary. These techniques can be categorised into two approaches, which utilise both bottom-up information and a top-down analysis (Williams, 2010, 2009). The bottom-up approach uses expenditure and revenue information sent by all spending units and revenue units to cash managers in the form of Disbursement Plans and Revenue Projections. The top-down approach relies on historical data of actual disbursements and revenues stored in the cash manager's database. It analyses aggregate spending and revenue by all government agencies over time to forecast future cash flows. This study aims to develop a government cash forecasting model with accuracy that meets an acceptable level of materiality for the cash manager. It is almost impossible to have an exact prediction. While most cash managers have their own preferences regarding the acceptable level of government cash forecasting accuracy, the ultimate goal is to have a model with the smallest margin of error.

Research focussing on the use and importantly the accuracy of government cash forecasting is rare. Storkey (2003), Williams (2009, 2004, 2010),) Mu (2006), and Lienert (2009, 2008), all state that cash flow forecasting research has been treated as complementary to the main topic of GCM. Storkey (2003) pointed out the advantages of using technology as a tool to develop effective cash management and forecasting systems. Williams (2009, 2004), proposed steps to enhance the cash management function, which included government forecasting ability as one of them. Williams (2009, 2010), also mentioned using bottom-up and top-down approaches to construct cash forecasting models. Mu (2006) emphasised the necessity of cash forecasting systems as one of the building blocks of effective GCM. Lienert (2009) extended the work of Williams (2009, 2004), on the sequencing of GCM reforms by introducing key features of a cash forecasting framework. Moreover, Mu (2006) provided information for cash managers to strengthen government cash forecasting ability. A discussion on how to develop forecasting models in various settings is provided in the next section.

2.5. Previous Studies on Forecasting Modelling

In general, most current studies on predictive modelling are developed following three main modelling approaches: statistical (e.g. ARIMA-based method), machine learning (e.g. ANN-based method), and hybrid technique which combines both statistical and machine learning approaches. Some studies focus on one particular method to construct forecasting models such as Huang (2019), Meier et al. (2018), and Alkhazaleh (2018) while others (e.g. Alam et al. (2018), Usmani et al. (2018), and Ülke et al. (2018)) provide a comparison between the available techniques in order to establish a model with the best accuracy. Table 2.2. lists some of the recent studies on developing forecasting models across five discipline areas.

Table 2.2: List of current studies on forecasting models

Method	ARIMA	ANN	Hybrid
Economics	• Huang (2019) • Adubisi et al. (2018) • Nyoni (2018) • Abdulrahman et al. (2018)	• Ülke et al. (2018) • Thakur et al. (2016) • Cogoljević et al. (2018) • Ristolainen (2018)	• Yan and Zhao (2018) • Inthachot et al. (2016) • Ince and Trafalis (2017) • Ghasemiyeh et al. (2017)
Finance	• Afeef et al. (2018) • Dong et al. (2017) • Alkhazaleh (2018)	• Jang and Lee (2018) • Ding et al. (2015) • Moghaddam et al. (2016) • Hiransha et al. (2018)	• Weerathunga and Silva (2018) • Khashei and Hajirahimi (2018) • Usmani et al. (2018)
Agriculture	• Khan et al. (2015) • Ilić et al. (2016) • Ali et al. (2015)	• Salman et al. (2018) • Zhang et al. (2018a) • Rathod et al. (2018)	• Alam et al. (2018) • Wang et al. (2018) • Rahim et al. (2018)
Transportation	• Kumar and Vanajakshi (2015) • Yan et al. (2017) • Miller (2018)	• Sharma et al. (2018) • Bartlett et al. (2018) • Ji and Hou (2017) • Zhao et al. (2017)	• Sulistyowati et al. (2018) • Hosseini and Shabanian (2018) • Li et al. (2017)
Power system	• de Oliveira and Oliveira (2018) • Yukseltan et al. (2017) • Yatiyana et al. (2017)	• Chakravorty et al. (2018) • Yadav et al. (2018) • Meier et al. (2018) • Li et al. (2018)	• Zhang et al. (2018b) • Chinnathambi et al. (2018) • Camelo et al. (2018)

Furthermore, regarding its nature, government cash forecasting can be seen as a cash demand prediction in the public sector. However, studies on predicting cash demand are dominated by research in the private sector area. Therefore, it is reasonable to argue that the modelling procedure undertaken in the private sector is also applicable to the public sector. On the other hand, the previous literature reveals that the most common case of cash demand forecasting studies in a private sector setting is the ATM cash demand forecasting modelling. Despite its different background, cash demand prediction in the public sector (e.g. government cash forecasting) and in the private sector (e.g. ATM cash demand forecasting) share the same research characteristics as summarised in Table 2.3.

Table 2.3: Characteristics of cash demand prediction in the public and private sectors

	Private sector (e.g. ATM cash demand forecasting)	**Public sector** (e.g. government cash forecasting)
Subject	Customers' money	Spending units' budget allocations
Purpose	Providing cash in the ATM ready for customers for withdrawal in the most cost-effective ways.	Providing cash ready for spending units to fulfil government obligations and deliver public services in the most cost-effective ways.
The timing	On customers' preferences	On spending units' preferences
The quantity	On customers' preferences	On spending units' preferences
Consequences of Overstocking	Opportunity cost	Opportunity cost
Consequences of Understocking	Customer dissatisfaction (bank's credibility lost)	Failure to fulfil government obligations and deliver public services (government's credibility lost)

Many studies on cash forecasting, such as those of Arabani and Komleh (2018), Bhandari and Gill (2016), Nemeshaev and Tsyganov (2016), Catal et al. (2015), and Dandekar and Ranade (2015), were conducted using various techniques with the purpose of finding the most accurate forecasting model. Those techniques included statistical approaches and machine learning methods. Some of the research compares several methods to determine the best models of cash forecasting.

Arabani and Komleh (2018) used statistical and machine learning methods to predict the ATM cash demand for Iran banking. The results confirm the superiority of the forecasting model developed based on the machine learning method over the one that is constructed using the regression model. Moreover, the researchers argued for the use of deep learning neural networks as the preferred machine learning method.

Bhandari and Gill (2016) incorporated calendar effects on ATM cash demand in modelling the future cash requirements. The FFNN technique of the ANN method was used to develop the ATM cash demand forecasting model. The result suggested that it is possible to use calendar effects as predictors for a cash demand projection problem.

Furthermore, a study by Nemeshaev and Tsyganov (2016) focussed its discussion on the pre-processing phase of the input data when building a forecasting model for ATM cash withdrawals. The research provides an evidential argument for the use of the machine learning method in facilitating a reasonable accuracy of the ATM daily cash withdrawals. Nemeshaev and Tsyganov (2016) also stressed the importance of having the appropriate number of training datasets in the successful development of the forecasting model using the machine learning method.

Catal et al. (2015) argued that the accuracy of a forecasting model can be enhanced by investigating the calendar effect on the data. In the case of the study, the performance of the prediction model of the ATM cash demand increases when considering the characteristics of cash demand on particular days. For example, the ATM cash withdrawals in the weekends are higher than on the weekdays, therefore, the binary number of 1 and 0 are set to define weekday and weekend cash demand respectively. The result demonstrates the superiority of the forecasting model that utilised the statistical method over the one that was developed using the ANN-based technique.

Similarly to Catal et al. (2015), Dandekar and Ranade (2015) took the calendar features of the historical data (e.g. payday, weekdays, and weekend) as a determinant factors in achieving a reasonable forecasting accuracy of the ATM daily cash demand. The research also used both statistical and machine learning techniques to establish a forecasting model. However, in contrast to the finding proposed by Catal et al. (2015), the result suggested that developing a forecasting model based on machine learning technique is better than the one with the statistical method.

Venkatesh et al. (2014) clustered the ATM cash demand patterns based on the calendar effect, i.e. day-of-the-week, in order to achieve a reasonable accuracy of the forecasting model. Multiple ANN-based methods implemented to each cluster separately for prediction modelling purposes. The results showed that the cash demand projections were improving while considering the calendar effect on the ATM cash demand patterns compared to none.

Some works have been conducted as an effort to develop a cash forecasting model in the public sector setting. Sumando et al. (2018) utilised monthly data of government expenditure grouped by type of expenditure. An independent cash forecasting model was built using multiple statistical-based methods for each group. The results showed that the best method to develop a government cash forecasting is inconclusive. Each type of expenditure proposed a different method to be the best procedure to develop a government cash forecasting model.

Iskandar et al. (2018) employed statistical, machine learning, and hybrid techniques to predict weekly government cash requirements. Contrary to Sumando et al. (2018), the research focussed only on the type of expenditure that is considered as intermittent expenditures in developing a government cash forecasting model. Iskandar et al. (2018) classified an expenditure as an intermittent expenditure when the timing and amount of government expenditure are not predetermined, by its nature. Opposite to intermittent expenditures, some routine expenditures are relatively fixed in terms of their timing and amount. Personnel expenditure is always due on the predetermined payroll date with an explicit amount of cash needed, for instance. Moreover, Iskandar et al. (2018) were considering goods expenditures, capital expenditures, and social aid expenditures as intermittent expenditures. Apart from proposing one specific model for each type of intermittent expenditure, the study also constructed a forecasting model utilising total cumulative intermittent expenditures. However, the finding remains vague as each type of intermittent expenditure suggested a different method as the best technique to produce a reliable government cash forecasting model. Therefore, Iskandar et al. (2018) argued that the best procedure to build a government cash forecasting model is subject to the data and the performance evaluation tools employed.

The preceding studies on predicting the future cash demand, both in public and private sectors, failed to conclude the best procedure to develop a reliable forecasting model. Pursuing the same purpose, this study takes advantage of the literature on time series forecasting modelling to design a framework for the conducted experiments.

Following Iskandar et al. (2018), this study focusses on government intermittent expenditures to develop a government cash forecasting model. As presented in Bhandari and Gill (2016), Catal et al. (2015), Dandekar and Ranade (2015), and Venkatesh et al. (2014), the present study also considered calendar effects on government intermittent expenditures as predictors. Moreover, in order to ensure that the best procedure on developing a government cash forecasting model is achieved, the present study utilised most methods proposed by the previous study as presented in Table 2.2. which were laid on statistical, machine learning, and hybrid model approaches. Supporting Arabani and Komleh (2018), this study employs the state-of-the-art in machine learning method which is a deep learning neural network technique to build a government cash forecasting.

2.6. Summary

The purpose of this chapter is to provide the relevant literature to the presented topic in this study. As has been emphasised throughout this chapter, having a government cash forecasting model with accuracy that meets an acceptable level of materiality for the cash manager is crucial not only for the government but also for the whole economy and beyond. The Greek crisis in 2010 is an example of where the burden caused by unanticipated borrowing to cover the government spending led to a shock in the Greek economy with the possibility of a contagion effect to other economies in the region.

In the area of PEM, the budget is seen as an instrument for the government to influence the economy while the ultimate objectives of the PEM are to maintain aggregate fiscal discipline, to allocate resources in accord with government priorities and to promote the efficient delivery of services. Both can be done through the three aspects of PEM: budget preparation, budget execution, and cash management. One of the main features of PEM is that effective government cash management is essential to maintain the availability of government money for public services while not causing unsustainable spending which provides risks to long term economic stability.

This chapter has discussed various means to achieve effective government cash management. It has highlighted the ability to predict government cash needs in the near future as a central feature of effective government cash management. Despite its importance, the quality of government cash forecasting is poor, especially in developing counties. To strengthen the accuracy of cash forecasting, the literature indicates that the cash manager should establish a database to collect historical data of inflows and outflows and analyse patterns of cash flows in order to develop an accurate cash forecasting model. It is also necessary to observe time constrained events and information that affects cash flows (e.g. public holidays, payment dates on weekends, and payment lead times) and making regular adjustments to the forecasting model to ensure its accuracy.

Studies focussing on developing forecasting models have been conducted in various settings as noted in Table 2.2 and in the private sector as noted in Table 2.3. Overall, models were built utilising statistical, machine learning and hybrid methods. There is relatively little research focussed on developing government cash forecasting models. Therefore, an aim of this study is to investigate the best ways to construct a government cash forecasting model with accuracy that meets an acceptable level of materiality for the cash manager in a public sector setting. In order to accomplish this aim, this study uses the Indonesian government as the case. An overview of government cash management in Indonesia is provided in the following chapter, together with a review of the existing government cash forecasting system that is used by the government cash manager in Indonesia. The best practice of government cash management discussed in this chapter is used to review the implementation in Indonesia.

Chapter 3 Government Cash Management and Forecasting in Indonesia

3.1. Introduction

The role of an effective GCM in achieving a sustainable economy was detailed in Chapter 2. The chapter also provides a guideline for the government cash manager to attain a better quality of GCM by introducing what is considered in the literature to be 'best practice' in effective government cash management. One of the main features is enhancing the accuracy of short-term government cash requirement predictions which can be done by establishing a reliable government cash forecasting model.

In this chapter, GCM and government cash forecasting practice in Indonesia are overviewed according to the best practice explained in Chapter 2. This chapter is intended to provide some background information on GCM and the existing government cash forecasting system, along with the characteristics of government expenditure, in Indonesia. Therefore, the description given in this chapter gives a firm practical foundation for this study.

An overview of GCM reform in Indonesia is provided in Section 3.2, while more detail on existing government cash forecasting arrangements in Indonesia is presented in Section 3.3. Section 3.4. details the characteristics of expenditure in Indonesia. Lastly, the summary for this chapter is presented in Section 3.5.

3.2. Government Cash Management Reform

The 1997 Asian financial crisis triggered public expenditure management reform in Indonesia. The crisis disclosed the institutional and structural weaknesses of public management in most Southeast Asian countries including Indonesia (Widodo et al., 2014). It forced the Central Bank of Indonesia to change its exchange rate policy from manage floating to independently floating regime on 14 August 1997 (Hernandez & Montiel, 2003). Such arrangement led to a sharp depreciation on the exchange rate.

The Indonesian Rupiah (IDR) reached the lowest point of 16,650 against the U.S. dollar in mid-June 1998, equal to 85% of cumulative depreciation since June 1997 (Lane et al., 1999). With the primary budget deficit financing sourced from foreign debt, the government burden inflated. According to the World Bank data, the ratio of Indonesian government debt to the gross domestic product (GDP) jumped from less than 25% in 1996 to more than 50% in the following years. The government's failure to anticipate the cost of borrowing caused by the depreciation of IDR created fiscal instability (Widodo et al., 2014). In addition, the government's credibility was questioned when many macroeconomic and financial indicators were off their target.

This situation was complicated by a rapid transition from a centralised state to a decentralised system of government in 2001. It transferred much of the responsibility for public spending and services delivery from the central government to over 400 local governments (Widodo et al., 2014). As a consequence of this policy, the central government had to provide the transfer fund for all those local governments. Whilst the government resources are limited, the cash management function in Indonesia needs to be advanced to minimise the cash manager's adversity. In response to such conditions, a gradual public expenditure improvement was introduced in 2003.

Cash management reforms were a backbone of the public expenditure management reforms in Indonesia. The first part of the reform was the enactment of a legal and regulatory treasury framework in 2003 and 2004 which focussed on modernising government cash management[1]. The laws give specific objectives for the government cash manager in managing the government money, which are: (1) securing the readiness of money to fund the government obligations; (2) optimising returns from cash surpluses or borrowing some amount of cash to deal with a shortfall in the most effective and efficient ways; (3) providing cash to spending units in line with their disbursement plans to meet their dedicated governmental programmes; and (4) ensuring payments to spending units' suppliers are executed in a timely manner (Widodo et al., 2014).

[1] Public Finance Act 2003, Public Treasury Act 2004, and Audit of Public Finance Act 2004.

The Public Finance Act 2003 mandated the Minister of Finance as Chief Financial Officer to be in charge of government cash management. The Minister of Finance appointed the Directorate General (DG) of Treasury – a unit Echelon I under the Ministry of Finance – to become the government cash manager responsible for managing all government money (Minister of Finance, 2004). As a cash management unit, the DG of Treasury also became responsible for reforming the GCM in Indonesia.

As described in Chapter 2, there are four stages in developing more effective government cash management: establishing the TSA; developing a cash forecasting system; "rough tuning" the government cash management; and "fine tuning" the government cash management. One of the critical components of an effective GCM is the establishment of TSA (Mu, 2006, Lienert, 2009, Williams, 2010). Its urgency is escalating with the lack of an advanced government account system. In such case, the first priorities for the GCM reform agenda should be implementing the TSA (Pattanayak & Fainboim, 2010). Under the TSA rules, all government cash balances have to be consolidated into a single account, preferably at the central bank (Williams, 2010). In the Indonesian context, the TSA is held by Bank Indonesia (BI), the central bank of the Republic of Indonesia.

Implementation of TSA in Indonesia was a gradual process. Prior to the TSA, the cash manager authorised the local treasury offices (LTOs) to establish multiple accounts at commercial banks to manage inflows and outflows of government money. The LTOs opened operational accounts, to accommodate payments based on the type of expenditure made by the spending units. For each account, the LTOs kept a certain amount of money as the minimum cash balance. On the revenue side, the cash manager made an agreement with the commercial banks to open revenue collecting accounts for the purpose of collection of the government revenues. Based on this contract, the commercial banks were allowed to keep the revenue for a couple of days before transferring it to the cash manager's account at BI. The cash balances for both operational and revenue collecting accounts were unremunerated. In total, such accounts were numbering in the thousands (Widodo et al., 2014).

In order to deliver their obligations, the spending units were allowed to have petty cash for covering their expenditures on small items and to collect revenues, based on the cash manager authorisation, from stockholders. As a consequence, it was permissible for spending units to have an account at the commercial bank for the purpose of the petty cash accounts and the revenues accounts. However, the cash manager did not have any control over the opening of the accounts. There was no regulation mandating the spending units to report the accounts to the cash manager. Therefore, the number of accounts was unrecorded. Such arrangements hindered the cash manager to optimise the cash balances accrued from the spending units' accounts (Widodo et al., 2014).

The formation of TSA in Indonesia started with verifying more than tens of thousands of government bank accounts operated by the LTOs and spending units across the country (Widodo et al., 2014). A dedicated task force was employed to collect and investigate the data of all government accounts which included the identification and verification regarding the ownership of the accounts and the legal basis for the establishment of the accounts. In its conclusion, the team decided whether to keep or to close the accounts. The remaining balance from the terminated accounts was then transferred to the TSA (Minister of Finance, 2007a).

Since the clarification of the government accounts, the LTOs have only been permitted to open the operational and revenue collecting accounts at a commercial bank that is appointed by the cash manager. This restriction allows the cash manager to make some arrangements with the commercial banks regarding the implementation of TSA, especially the zero balance accounts arrangement (Widodo et al., 2014). Zero balance accounts are set up to all the LTOs accounts, so the residual balances are transferred back to the TSA daily. TSA implementation for expenditure accounts began with some piloting for spending accounts in 2006 (Minister of Finance, 2006a), followed by full implementation in 2007 (Minister of Finance, 2007b). On the other hand, TSA implementation for revenue accounts was piloted in 2009 (Minister of Finance, 2009b) and comprehensively employed in 2010 (Minister of Finance, 2010b). The new arrangements were also introduced to the government bank accounts held by the spending units. A formal consent from the cash manager is required for the spending units to open new government accounts. A set of regulations has been enacted to facilitate the cash manager's control over the accounts (DG of Treasury, 2007). With all government accounts under its monitoring and control, the cash manager has the flexibility that is needed to manage the government money effectively.

As a comprehensive measure, the government cash forecasting reform in Indonesia also includes an improvement in government accounting and payment operations. IFMIS is the tool that is commonly used to aim for these goals. IFMIS refers to an integrated system of public expenditure management processes for financial management of the line ministries and other spending agencies. It includes budget formulation, budget execution, and accounting (Diamond & Khemani, 2005). An IFMIS stores not only the information on expenditure but also other information such as the budgets, cash flows, assets and liabilities. (Rodin-Brown, 2008).

IFMIS in Indonesia is a long ongoing process started in 2004 when the Government Financial Management and Revenue Administration Project, sponsored by the World Bank, was initiated. The project adopted an IFMIS framework with an aim to modernise Indonesian public expenditure management (Minister of Finance, 2011). After completing the preparation phase, a series of piloting and rolling out of IFMIS was employed in 2013 (Minister of Finance, 2013) and fully implemented in 2015 (Minister of Finance, 2014a). An IFMIS generally consists of several modules that perform different functions (Rodin-Brown, 2008). For the Indonesian context, IFMIS contains a budgeting module, commitment module, payment module, cash management module, revenue module, and accounting and reporting module.

Solid coordination amongst government agencies is important for effective cash management. It assists the cash manager to provide accurate projections of short-term cash flows and a foundation for more active cash management. Information sharing between the cash manager, spending units, and revenue-collecting units will ensure the future cash flows are known in a timely manner (Williams, 2010). To enhance the interaction with the cash manager in the GCM framework, spending units in Indonesia are provided with an integrated financial application software called Spending Units' Financial Application System (SUFAS) developed by the cash manager. This new software integrates all the presently existing independent computer-based applications used by the spending units and functions as a gateway application for IFMIS. Once the IFMIS can access the information from SUFAS, real-time and detailed financial information will be accessible for the cash manager through the IFMIS cash management module which is valuable for cash forecasting and other cash management purposes (Widodo et al., 2014). SUFAS has been piloted since the budget year 2016 began (Minister of Finance, 2015b).

Information technology plays a significant role in enhancing the quality of GCM (Williams, 2010; Lienert, 2009). It provides valuable data for the cash manager to develop a reliable cash forecasting system. Accurate information regarding the cash needed by the spending units in the future needs to be passed on to the cash manager in the very first instance. Hence, an appropriate amount of money can be presented in the most cost-effective ways. Although the IFMIS has an ability to deliver comprehensive financial information amongst the stakeholders (Khemani & Diamond, 2005), Williams (2010) argued that the establishment of IFMIS is not a must when a reliable cash forecasting system is the objective. A separate information network and database are reasonable to develop for short-term cash flow projections purposes as long as they are fully controlled by the cash manager (Williams, 2010).

In general, the international best practice on developing a government cash forecasting system tends to follow bottom-up and top-down approaches. While the bottom-up approach relies on the information reported by the spending units to the cash manager, the top-down approach utilises historical data gathered from a database held by the cash manager (Williams, 2009, 2010). Minister of Finance Regulations number 192/PMK.05/2009 regulated the government cash forecasting system in Indonesia. The system adopted a bottom-up approach when developing a government cash forecasting model. However, a report revealed by Widodo et al. (2014) stated that the accuracy of the cash forecasting model was still below the expectations of the cash manager. Therefore, a new mechanism was introduced with the aim to increase the accuracy of the forecasting model (Minister of Finance, 2014b). More detailed discussion on the government cash forecasting system in Indonesia is presented in a dedicated section below.

With the establishment of the TSA and the government forecasting system in Indonesia, the next phase is to improve the basic arrangements from previous stages. Williams (2009) divided the next episode into “rough tuning” and “fine tuning” improvements. In the “rough tuning” stage, a minimum cash balance is set and maintained in manageable ways. One of these is by minimising the carrying cost of money. This can be done by remunerating the idle cash balance of government accounts. In this regard, a memorandum of understanding on the coordination of GCM between the Minister of Finance and the Governor of BI was enacted in 2009 (Widodo et al., 2014). This agreement stipulated the BI was to pay remuneration for all government cash balances kept in the central bank. The amount of remunerations is varied depending on the accounts where the money is located. The interest rate on the TSA is 0.1% per annum of daily cash balances. The same memorandum of understanding also mandated the BI to pay remuneration to the cash manager with the interest rate at 65% per annum of Bank Indonesia policy rate for IDR accounts, 65% per annum of the Fed Funds rate for USD accounts, and 65% per annum of the reference rate of the home currency for other currencies accounts (Minister of Finance, 2009d). The arrangement is a mutually beneficial situation. From the cash manager’s point of view, BI delivers a zero-risk investment for government money with a benefit of remuneration for government revenue. On the other hand, the retention of government money in BI eliminates the cost regarding sterilising the excess liquidity caused by government cash placement in commercial banks, which is advantageous for the BI (Widodo et al., 2014). In addition, the memorandum of understanding set the daily minimum cash balance for all rupiah accounts at IDR 2,000,000,000,000.00 (two trillion rupiahs) and equivalence of USD1,000,000.00 (one million United States dollars) for foreign exchange USD and non-USD accounts (Minister of Finance, 2009d). Such an amount of money is intended as a reserve for unexpected government expenditures (Minister of Finance, 2010a).

Similar to the cash manager's accounts, remuneration also gained from the cash balance of government accounts is held by the spending units. The Indonesian cash manager made an agreement with the commercial bank on implementing the Treasury Notional Pooling (TNP) covering expenditure accounts (Minister of Finance, 2009a) and revenue accounts held by spending units (Minister of Finance, 2009e). TNP is a balance consolidation management program that allows performing the consolidation without doing any overbooking or cash transfer for every revenue/expenditure spending unit's account at the end of the day. This arrangement is possible due to the advance of "virtual" consolidation. By "virtually pooling" the daily residual balances of the spending units' accounts, the cash manager receives remuneration on idle government cash balances (Widodo et al., 2014).

Another aspect of GCM to be improved in the "rough tuning" is the cash and debt management coordination. While having one integrated institution responsible for cash and debt management function is preferable, Williams (2010) stressed close coordination between both functions as the key for an effective GCM. In the case where both functions are run by two separate units, the establishment of forums that allow interaction at policy and operational levels are critical (Williams, 2010).

A new institutional arrangement was introduced in 2006 by the Indonesian government. The Minister of Finance, as chief financial officer of the government of Indonesia, separated the debt management function from the cash management function. The responsibilities for searching the source of financing and managing the debt portfolio were handed over to the DG of Debt Management while the government cash management function remains with the DG of Treasury (Minister of Finance, 2006b). To make sure debt strategy is in line with cash management and other policies, the Ministry of Finance formed an Asset and Liability Management Committee (ALMC) and Cash Planning Information Network (CPIN) (Widodo et al., 2014).

As a policy level coordination, the Minister of Finance chairs the ALMC with the Vice Minister as the Deputy Chair. The ALMC meeting is held at least once a month or more frequently at the request of the Minister of Finance. The ALMC may invite other participants when required. The ALMC meeting mainly discusses topics presented by its members. The topics are as follows: (1) the economic and market outlook presented by the Fiscal Policy Office (FPO), DG of Treasury, and DG of Debt Management; (2) the revenue projections presented by FPO, DG of Tax, DG of Customs and Excise, and DG of Budget; (3) the expenditure projections presented by DG of Budget and DG of Fiscal Balance; (4) the cash projections and financing presented by DG of Treasury and DG of Debt Management; and (5) the assessment of risk to the balance sheet and budget activities, assessment of the future cash requirements and new financing, along with recommendations on changes in policy, if required (Widodo et al., 2014).

At the operational level, CPIN is represented by technical staff from various Echelon I of the Ministry of Finance such as DG of Budget, DG of Treasury, DG of Debt Management, FPO, and others. CPIN holds a discussion at least once a month and releases a monthly cash forecasting report for the Ministry of Finance. This report is prepared using the spending units' historical data of cash flow, along with assumptions about key macroeconomic and monetary indicators (Widodo et al., 2014).

The Indonesian government is committed to advance its continuous improvement in achieving an effective GCM. As mentioned by Lienert (2009), the last sequence in realising an effective GCM is "fine tuning" the arrangements of GCM itself by shifting the cash manager's focus towards more active daily management. It is policy-intensive and requires detailed agreements with other policymakers and the banking sector due to its direct intervention in the money market (Pessoa & Williams, 2012). Lack of coordination amongst the related parties increases the risk of economic instability caused by destructive contradicting policy (Williams, 2010).

A regulation framework regarding how to handle government cash surplus and/or deficit has been issued by the Indonesian government since 2010. In this regulation, cash surplus happens when daily cash balance is exceeding the minimum daily reserve balance target which is IDR 2 trillion. On the contrary, the cash deficit is defined as a shortfall in the daily cash balance compared to the minimum daily reserve balance target. A number of instruments are set for placement of surplus or deficit financing which include placement or withdrawal of state cash in the central bank and the commercial banks, purchase or sale of Government bonds from the secondary market, and Repo or reverse Repo (Minister of Finance, 2010a). Stating its dedication for active daily cash management policy, the Indonesian government improved the rules and procedures for the placement of surplus and the financing of the deficit in 2016 (Minister of Finance, 2016).

The active daily cash management policy of the cash manager introduces the concept of the Treasury Dealing Room (TDR). The TDR allows the cash manager to manage cash liquidity through money market instruments in response to cash surplus and/or deficit (Widodo et al., 2014). The TDR differs from the DG Debt Management Dealing Room in term of time horizon of its instrument to be used. Whilst the DG Debt Management Dealing Room focusses on the long-term instrument, the TDR concentrates on the short-term instrument (Widodo et al., 2014). The TDR will allow the cash manager not only to use the government cash surplus for cash placement in the central bank and commercial banks, purchase of government bonds from the secondary market, and reverse Repo (Minister of Finance, 2010a) but also other money market instruments (Widodo et al., 2014). The differences between the TDR and the DG Debt Management Dealing Room are summarised in Table 3.1.

Table 3.1: The differences between the TDR and the DG Debt Management Dealing Room

	DG Debt Management Dealing Room	Treasury Dealing Room
Aim	Issuance and redemption of the Government securities	Manage cash liquidity through money market instruments
Instruments to be traded	Debt instruments, such as: T-Bonds; T-Bills; State Sharia Securities	Money market instruments, such as: cash placement in the central bank and the commercial banks; purchase of government bonds from the secondary market; Reverse Repo
Horizon	Long term (90 days or more)	Short term (maximum 90 days)

Source: Widodo et al. (2014)

The Indonesian cash management reform in Indonesia is considered as a success story (Widodo et al., 2014). However, providing a reliable government cash forecasting model with accuracy that meets an acceptable level of materiality for the cash manager is among the challenging list to be overcome. The next section discusses the Indonesian government's effort to develop a government cash forecasting model.

3.3. Government Cash Forecasting

A valuable lesson to learn from the Greek crisis in 2010 is that the failure of the cash manager to effectively manage the government money not only disrupts the domestic economy but also threatens a wider community. In the case of the Greek crisis, the unforeseen cost of borrowing, to cover government expenditures, became an economic burden for Greece with the possibility of a contagion effect to other economies in the region (Arghyrou & Tsoukalas, 2011). Therefore, having an effective GCM is crucial for achieving a sustainable economy.

Many studies have suggested the development of a government cash forecasting model as a fundamental element of an effective GCM. Mu (2006) developed building blocks of an effective GCM with the government cash flow forecasting as one of the pillars. Lienert (2009) and Williams (2010) considered the ability to accurately predict short-term cash flow as one of the basic requirements to achieve effective GCM. A detailed explanation regarding the importance of the development of government cash forecasting model has been presented in Chapter 2.

As part of the GCM reform, developing a government cash forecasting model has become the central point of the Indonesian government concern. The cash manager is mandated by the Indonesian government to administer government cash planning as the underlying information for establishing the GCM strategy on handling government cash surplus or shortfall (Government of Indonesia, 2007). From the perspective of Williams (2010), the development of government cash forecasting in Indonesia adopts both bottom-up and top-down approaches.

Following the bottom-up approach, all spending units in Indonesia are required to submit a regular disbursement plan to the cash managers, along with its updated reports. The reports are made on a monthly, weekly, and daily basis. Every document has different reporting rules depending on its reporting periods and type of expenditure. The cash manager uses the report as input data to establish a government cash forecast (Minister of Finance, 2009c). On the other hand, an inter-directorate committee in the Ministry of Finance – CPIN – prepares a top-down monthly cash prediction report based on historical data of revenues along with the assumptions of key macroeconomic and monetary indicators. The main focus of the report is to update the monthly cash forecasting for particular mandated expenditures to be executed by the Ministry of Finance such as interest payments, subsidies, and other expenditures. Therefore, the use of the top-down approach in the government cash forecasting system is a complement to the bottom-up approach (Widodo et al., 2014).

At the initial stage, spending units prepare the disbursement plans for the respective budget year, on a monthly basis, and include it in their Budget Execution Document. Such disbursement plans are a rough estimation for which spending units need to make regular updates throughout the year. Such an estimation is useful for the cash manager to set a preliminary GCM strategy for the whole year at the beginning of the budget year. However, most of the spending units see the plans as a formality measure rather than a practical exercise. Instead of carefully planning their provisional expenditure in the future, the spending units distribute the annual budget allotment equally for each month along the budget year (Widodo et al., 2014). Therefore, the information held by the cash manager becomes meaningless.

To overcome the problem, a disbursement plan reporting mechanism has been introduced by the cash manager. It comprises monthly, weekly, and daily reporting mechanisms. Similar to the disbursement plan that is attached to the spending units' Budget Execution Document, the monthly disbursement plans report the proposed spending unit's expenditure for each month of the budget year. The monthly disbursement plan needs to be sent to the cash manager by the tenth working day after receiving the budget execution document. The monthly expenditure allocated in the disbursement plans is the ceiling for the respective month. However, the ceiling is not rigid. There is an update mechanism that allows the spending units to carry over the remaining budget allocation on a particular month into the following months (Minister of Finance, 2009c).

Moreover, the monthly disbursement plans are detailed into a four-week period for each respective month based on the calendar date such that the first seven days of the month are considered as the first week and so forth. This weekly disbursement plan is a bimonthly report that needs to be sent to the cash manager before the fifth working day of the beginning planned week (Minister of Finance, 2009c). Furthermore, for day to day forecasting purposes, the spending units are required to break down their weekly disbursement plans for each day of the week into daily disbursement plans. Every spending unit has to prepare the report every week and send it before the second working day of the beginning planned day. The same updating mechanism applies to ensure the remaining budget allocation on a particular reporting period is carried over to the succeeding periods accordingly (Minister of Finance, 2009c).

It is permissible for the spending units to adjust their reported disbursement plans by sending an update to the cash manager. The updated version of the disbursement plans needs to be received within the specified timeframes. The deadlines are before the third working day of the respective month for the monthly report, before the second working day of the respective week for the weekly report, and one working day before the respective day for the daily report. The disbursement plan and its updates are sent to the cash manager via the treasury office (Minister of Finance, 2009c). The timeframes for the disbursement plans reporting, along with their updates are summarised in Table 3.2.

Table 3.2: The disbursement plans reporting timeframes

Type of reporting	Reporting deadline	Updating deadline
Monthly	10 working days after receiving the budget execution document	3 working days before the updated period
Weekly	5 working days before the reported period	2 working days before the updated period
Daily	2 working days before the reported period	1 working day before the updated period

Source: Minister of Finance (2009c)

To improve government cash forecasting accuracy, a new arrangement was launched in 2014. The cash forecasting's new arrangement provides convenience for the spending units to regularly submit their disbursement plans by relaxing the reporting requirements. Under the new regulation, there are only two types of reporting periods which are monthly and daily. Very much like the previous procedure, the monthly disbursement plans report the projected spending unit's cash requirements for all type of expenditure. However, there is no need to update the monthly disbursement plans as long as the accumulation of the daily disbursement plans is not exceeding 20% of the monthly disbursement plans for a particular month (Minister of Finance, 2014b).

Another incentive is on the daily disbursement plans reporting. It simplifies the role of spending units in providing their daily disbursement plan to the cash manager by selecting a particular expenditure that needs to be reported in advance. While the previous policy made it mandatory for a spending unit to include all future transactions into the disbursement plan and its updates, the new regulation only requires the particular expenditure that is considered a “big transaction” to be reported to the cash manager. The arrangement was driven by the fact that approximately over 75% of the spending units in Indonesia is bestowed with a budget allotment of less than IDR 10 billion each comprising only 3.25% out of the total budget allotment (Widodo et al., 2014). Applying the Pareto principle, the cash manager is focussing only on the few spending units with large expenditure budget allocations for submitting their regular updated cash forecasting (Widodo et al., 2014).

The “big transaction” expenditure for each spending unit is varied. It depends on the type and location of treasury office with which the spending unit has been partnering. There are 178 treasury offices in Indonesia (DG of Treasury, 2013). For the sake of cash forecasting, the treasury offices are classified into three categories: (1) the type A1 treasury office that is located in the capital city of a province; (2) the type A1 treasury office that is not located in the capital city of a province; and (3) the type A2 treasury office, regardless of its location. The “big transaction” expenditure for each treasury office’s class is grouped into three classes regarding the value of the transaction itself. In total, there are nine groups of “big transaction” expenditure. The spending units are required to make the disbursement plan only for the transactions that match with one of the “big transaction” expenditure classifications (Minister of Finance, 2014b). The “big transaction” criteria and their respective timeframes are listed in Table 3.3.

The spending units report their daily disbursement plans a couple of days prior to the actual expenditure event. As presented in Table 3.3, all daily disbursement plans need to be received by the cash manager five days before the actual spending except for transactions A and B which are 15 days and 10 days before, respectively. Moreover, there is no obligation for the spending units to make an update on their daily disbursement plans except for transaction A which is 10 days before the actual expenditure and transaction B which is 5 days before the actual expenditure (Minister of Finance, 2014b).

Table 3.3: The list of the "big transactions"

Treasury Office	Type of Transaction	The value of transaction	Reporting Deadline
Type A1 located in capital city of the province	Transaction A	> IDR 1 Trillion	15 working days before
	Transaction B	> IDR 500 Million to IDR 1 Trillion	10 working days before
	Transaction C	> IDR 1 Million to IDR 500 Million	5 working days before
Type A1 not located in capital city of the province	Transaction D	> IDR 1 Million	5 working days before
	Transaction E	> IDR 750 Billion to IDR 1 Million	5 working days before
	Transaction F	> IDR 500 Billion to IDR 750 Billion	5 working days before
Type A2 regardless of its location	Transaction G	> IDR 500 Billion	5 working days before
	Transaction H	> IDR 350 Billion to IDR 500 Billion	5 working days before
	Transaction I	> IDR 200 Billion to IDR 350 Billion	5 working days before

Source: Minister of Finance (2014b)

Along with the policies on cash forecasting arranged by the cash manager, computer software has been released to facilitate spending units in preparing and sending their disbursement planning and its updates (Widodo et al., 2014). Such a facility is an effort to develop a reliable government cash forecasting model with accuracy that meets an acceptable level of materiality for the cash manager. However, despite the seminars and workshops on the regulations and technical measures of the disbursement plans, a review conducted by the cash manager discovered that no more than a half of the spending units were committed to submitting their cash projections according to the designed schedule. The review also showed that much of the planning was inaccurate. The report displayed that the performance of the government cash forecasting system is violating the cash manager's expectation. The onerous requirements for reporting the future expenditure, along with its updates, and the absence of sanctions to the noncompliance spending units are amongst the suspected sources of the problem (Widodo et al., 2014).

Furthermore, Widodo et al. (2014) argued that expanding the role of a top-down approach based on historical patterns might increase the accuracy of the expenditure prediction. Currently, the use of a top-down approach in the government cash forecasting system is limited to particular expenditure such as interest payments, subsidies, and other expenditures. This type of expenditure is authorised for the Ministry of Finance only. Utilising the top-down approach to develop a government cash forecasting model eliminates the requirements of the reporting from the spending units. Therefore, *with the fact described above,* the present study aims to develop a government cash forecasting model by following the top-down approach. In the next chapter, the characteristics of government expenditure in Indonesia are detailed.

3.4. Characteristics of Government Expenditure in Indonesia

Based on regulation, the Indonesian government classifies its expenditures into eight types of expenditures: personnel expenditures, goods expenditures, capital expenditures, interest, subsidies, grants, social aid expenditures, and other expenditures (Minister of Finance, 2015a). The spending units use personnel expenditures to pay the salary of government employees and other personnel compensations. Therefore, most of the personnel expenditures occur at the same time every month which is on payday. In order to function, the spending units require operational costs which are covered by goods expenditures. Such expenses are subject to the spending units' discretion regarding the timing and the amount of money required. To support their operational activities, the spending units occasionally involve the acquirement of new assets or the intensification of the assets' values. Capital expenditures are used to finance spending. Although the procurement procedures regulate the execution of the capital expenditures, it is the spending units' decision to make the timeframe for the completion of the project.

Interest expenditures are allocated to pay interest on the outstanding debts and other government debts costs. To maintain the purchasing power over some products, subsidies are given to state companies, government agencies, or other parties. Grants are the government transfers to other countries, international organisations, local governments, and communities. The cash manager coordinates and authorises spending units to execute interest, subsidies, and grants expenditures. Therefore, the disbursement of these types of expenditure is known to the cash manager prior to being performed.

In order to protect people or communities from the possibility of social risk and to maintain their welfare, the spending units are allowed to make a direct transfer of some amount of money, goods, or services. The outlays for such transfers are taken from social aid expenditures. It is up to the spending units' judgment to assess the potential social risk based on the duties and public services provided. Other expenditure includes reserve funding for natural disasters, social disasters, and other unforeseen events. By default, the delivery of other expenditure is planned based on particular events.

The characteristics of each type of expenditure affect the disbursement behaviour of spending units. By its nature, Indonesian government expenditure can be categorised into routine and intermittent expenditures. Routine expenditures, such as personnel expenditures, interests, subsidies, grants, and other expenditures, are the expenditures for which the time and the amount of funding are relatively easy to foretell. In most cases, the timing and funding for the routine expenditures are scheduled and relatively fixed. Some routine expenditures are known to the cash manager in advance since this is mandated by regulations.

On the other hand, the timing and the funding required for intermittent expenditures fluctuate over periods. This includes goods expenditures, capital expenditures, and social aid expenditures. The three expenditures are fully controlled and under the discretion of the spending units without intervention from the cash manager. For instance, goods expenditures are closely related to a spending unit's operational activities and only the spending unit itself knows when to execute the spending. Regarding the capital expenditures, the spending units are forced to synchronise their project with the procurement process. Delays at any stage of the process affect the spending units' disbursement plans. Some spending units' activities are reliant on certain factors for successful implementation. For example, agricultural projects depend on weather, soil condition, and any other natural features, therefore, it is purely on the spending units' decision regarding the best time to execute the project. The uncertainty of the intermittent expenditures makes this group of expenditures challenging to predict (Widodo et al., 2014).

It is mandated by the law 17/2003 on state finances that in order to be able to fulfil its obligations and deliver public services, the Indonesian government, together with its legislative body, establishes the state budget annually to start on 1 January and end on 31 December of the same year[2]. It compiles all potential revenues, planned expenditures, and possible financing. In practice, the annual state budget is broken down into budget execution documents for each spending unit. As annual financial planning, the expenditure displayed in the state budget is the maximum ceiling that every spending unit needs to obey (Government of Indonesia, 2004). Therefore, it is logical to consider that the disbursement of spending units is related to the availability of budget allotment for the respective expenditure. The maximum amount of expenditure of a particular spending unit for a specific day is subject to and will not exceed the remaining budget for that certain expenditure, for example.

Figure 3.1 shows the percentages of quarterly total intermittent expenditures for all spending units in Indonesia for the period of 2009 – 2015. As shown in Figure 3.1., the disbursement activity of spending units tends to be low in the first quarter of the budget year with an average of 7% approximately. The averages are increased in the second and third quarters with around 18% and 24% respectively. The increase continues in the last quarter of the budget year with an average of 51%. It is obvious from the sharply increasing trends shown by the graph that the spending units' expenditures are varied over time. The behaviour of spending units' disbursement tends to be low at the specific time of the budget year while rising high in the other parts of the budget year.

[2] Previously, the fiscal year in Indonesia was from 1st of April to 31st of March of the following year.

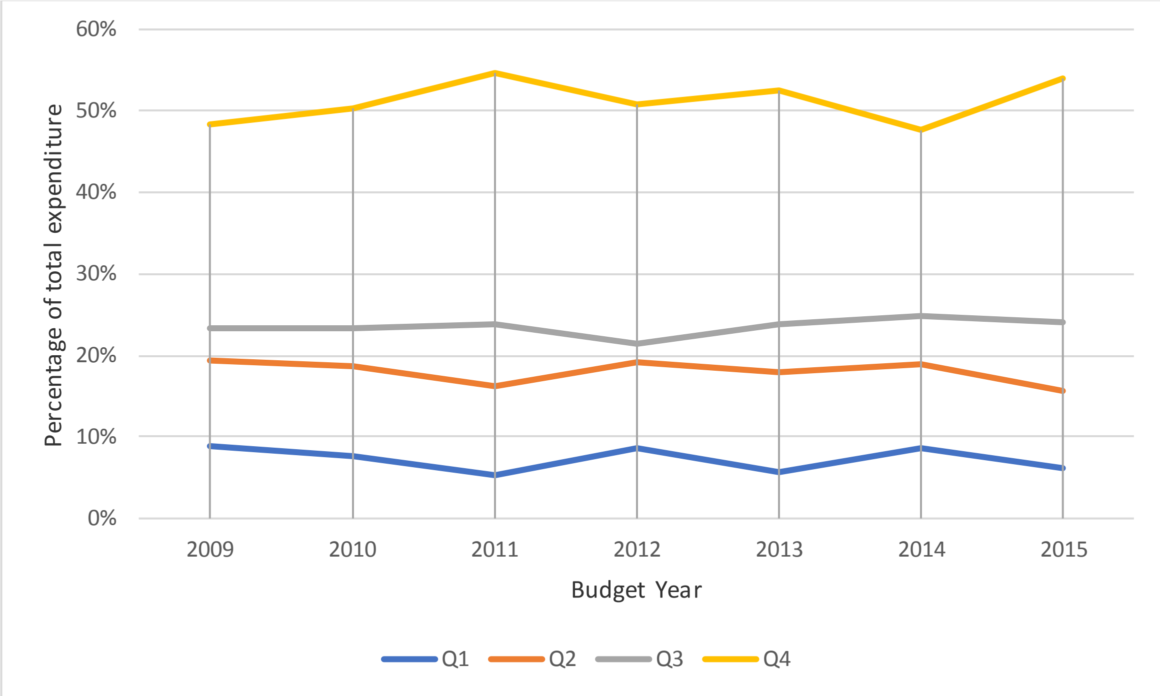


Figure 3.1: The proportion of quarterly disbursement pattern of total intermittent expenditure 2009-2019

Research conducted by Siswanto and Rahayu (2010) argued that in the context of the Indonesian government, an improvement on the budget policy affects the spending units' disbursement patterns. The research suggested that the escalation of the spending units' expenditures at the end of the budget year could be minimised by simplifying the administration requirements related to the budget execution document. Furthermore, since the introduction of the new fiscal year in 2001, spending units received the budget execution document as legal authorisation of their expenditure at the beginning of the budget year. In 2011, the Indonesian government established a new regulation on handing over the budget execution document to spending units. The new policy mandates the Ministry of Finance as CFO to present the document to the spending unit before the budget year begins. It was intended to encourage spending units to execute their programs early in the first half of the budget year such that the spending units' cash disbursement patterns are expected to smoothen (Minister of Finance, 2010c).

3.5. Summary

The purpose of this chapter was to provide understanding regarding the context of the Indonesian government for the present study. It discussed the GCM reform including the development of government cash forecasting in Indonesia. It also detailed the characteristics of government expenditure in Indonesia. This chapter revealed that although the Indonesian government has successfully implemented a GCM reform to enhance its effectiveness, providing a reliable government cash forecasting model with accuracy that meets an acceptable level of materiality for the cash manager is problematic.

Despite the success of implementing the best practice on guidelines to enhance the quality of GCM, the Indonesian cash manager is still facing difficulty when it comes to accurately predicting the future cash needed by the spending units in order to fulfil the government obligations and deliver public services. Improvements on the cash forecasting arrangements have been implemented to increase the accuracy of the model with a little accomplishment. Overall, the development of the cash forecasting model in Indonesia has adopted both bottom-up and top-down approaches. However, the use of the top-down technique has played a small role in the cash forecasting system. Therefore, the study is proposing the development of a government cash forecasting model following the top-down approach.

By regulation, the Indonesian government has classified eight types of expenditure: Employee Expenditures, Goods Expenditures, Capital Expenditures, Interest, Subsidies, Grants, Social Aid Expenditures, and Other Expenditures. Each of them is entitled to a specific characteristic that determines their disbursement behaviour. Moreover, the government expenditures in Indonesia can be categorised based on their disbursement behaviour into routine and intermittent expenditures. While forecasting the future routine expenditures is straightforward, the nature of intermittent expenditures makes the forecasting process challenging. As part of the state budget, the spending units' intermittent expenditures are subject to the ceiling stated on their budget execution document. The historical data show that the calendar effect plays a role in shaping the disbursement patterns of the spending units. Moreover, the literature suggests an enhancement in budget policy might contribute to the spending units' expenditure behaviours. For that reason, the present study has utilised the intermittent expenditures to construct a daily cash forecasting model by considering the property of the available funds, calendar effects, and policy implementation on the expenditures on the specific day. An in-depth discussion on the aim of the research and related questions that this study intends to address is detailed in the next chapter.

Chapter 4 Development of Research Aim and Questions

4.1. Introduction

The purpose of this chapter is to identify the key features from the literature, which will provide the basis of the key research aim and the subsequent development of the research questions posed in the study. The concept of PEM, GCM, and government cash forecasting was introduced in Chapter 2, together with the previous studies on a cash forecasting model. It provided the “best practice” for an effective GCM. Putting the idea into context for this study, Chapter 3 described GCM reform, the development of a government cash forecasting system, and the characteristics of government expenditure in Indonesia.

This chapter will draw a line to connect the concept of PEM, GCM, and government cash forecasting to its practice in the Indonesian government to provide a strong foundation for the research questions of the present study. A synthesis of the literature reviewed in Chapter 2 is provided in Section 4.2 below, followed by a discussion regarding how the outcomes of the literature review contribute to the achievement of the research aim and the development of the research questions in Section 4.3. A brief summary is provided in Section 4.4.

4.2. Synthesis of Literature

The structure of the literature review provided in this study reflects the funnel-shaped approach. In order to comprehend the importance of a government cash forecasting model in achieving an effective GCM, a deep understanding regarding positioning of GCM in the area of PEM is necessary. Therefore, the concept of PEM is first introduced before presenting the analysis on GCM as one of PEM’s aspects. Once the positioning of GCM has been made firm, the significance of having an effective GCM and how to achieve it are discussed. Lastly, the government cash forecasting model, as the main element of effective GCM, is elaborated.

As discussed in Chapter 2, the PEM is seen as the economic perspective of government expenditure. The economic theory of public spending is based on one of the two works of Adolph Wagner and John Maynard Keynes. While Adolph Wagner saw government spending as a response to the economic growth, John Maynard Keynes on the contrary, considered public expenditure as a trigger for economic development. However, the Greek crisis of 2010 shared a lesson to learn that regardless of which underlying theory a nation has followed, stable economic development depends on how the government manages its expenditures. Therefore, a good understanding of PEM and its aspects is necessary to have a sustainable economy.

In the discipline of PEM, the theme of discussions is the issue of allocating and using economic resources in order to fulfil the government's responsibilities and public services delivery. It comprises three critical aspects of PEM: budget preparation, budget execution, and GCM. Budget preparation generally deals with constructing the national budget following targeted macroeconomic indicators. During the budget execution, a number of procedures are set such that the implementation of the national budget meets the macroeconomic benchmarks. Nonetheless, the GCM focusses its attention to guarantee the national budget is funded in the most effective way in line with the macroeconomic framework.

Acknowledging the role of GCM in the economy, the literature described in Chapter 2 proposed guidelines on modernising GCM with an effective GCM as the ultimate goal. The guidelines involve four stages in developing a more effective GCM: (1) establishing the TSA; (2) developing a cash forecasting model; (3) "rough tuning"; and (4) "fine tuning". While the stages are a sequence in the implementation, the first two phases are the prerequisite for an effective GCM and the last two phases are required for the GCM improvement.

Since it is the precondition of an effective GCM, the development of a government cash forecasting model should be one of the main concerns for all nations in order to sustain their economy. However, not all countries are endowed with the luxury of having a reliable government cash forecasting model. Providing an acceptable accuracy of government cash forecasting model is problematic. Such a predicament is not necessarily a monopoly of the developing country, it is also suffered by a more developed country like the United Kingdom to a certain extent. Therefore, the present study proposes the procedure of developing a reliable government cash forecasting model with accuracy that meets an acceptable level of materiality for the cash manager.

There are two approaches on how to develop a government cash forecasting model: the bottom-up and top-down approaches. The bottom-up approach is where the budget executing agencies report their future expenditure plans to the cash manager. The reports are aggregated to form government cash forecasting. With this approach, the cash manager utilises the information gathered from the spending units to construct a government cash forecasting model. On the other hand, under the top-down regime, a government cash forecasting model is built employing the information that is known to the cash manager.

The information that is used to develop a government cash forecasting model following the top-down approach, such as historical data of spending units' expenditure, are kept in a database held by the cash manager. There is no requirement for the spending units to report their disbursement plans as happens with the bottom-up approach. It has a twofold advantage: (1) It reduces the time constraints on acquiring information for modelling purposes and (2) It relieves the spending units of the reporting requirements. Despite the advantages over the bottom-up approach, there are not many studies which focus on developing a government cash forecasting model using the top-down approach. Hence, the present study utilises the top-down approach to propose a government cash forecasting model.

Most research on the development of a forecasting model based on historical data in various disciplines has been carried out utilising various methods. These methods tend to centre on three main modelling approaches: statistical, machine learning, and a combination of both statistical and machine learning approaches called a hybrid model. Moreover, previous studies on developing a government cash forecasting model are rare. Some studies have employed monthly data of government expenditure, while weekly government spending data have been applied by others. Nevertheless, the studies have failed to provide consensus on the best method to develop a government cash forecasting model. The findings give the foundation for the research questions posed in the present study. The development of these questions is provided in the following section.

4.3. Development of Research Questions

As discussed in Chapter 2, having a reliable short-term government cash forecasting model is essential for effective GCM which, from a broader perspective, helps the government to achieve sustainable economic development. However, establishing a government cash forecasting model with accuracy that meets an acceptable level of materiality for the cash manager is challenging for most countries.

Furthermore, Williams (2010, 2009) stated that there are two approaches to developing a government cash forecasting model: bottom-up and top-down approaches. By its definition, this study synthesises the benefits of using the top-down approach over the bottom-up in developing a government cash forecasting model. Moreover, Mu (2006) specifically suggested developing a cash forecasting model from the historical data as one of the main strategies in strengthening the quality of government cash forecasting ability. Notwithstanding its superiority, the use of the top-down approach to develop a government cash forecasting model is still limited. Therefore, the research aim is: To develop a government cash forecasting model with accuracy that meets an acceptable level of materiality for the cash manager by utilising historical government expenditure data.

Based on previous studies on developing a government cash forecasting model, the performance of the model is subject to the data and method that are utilised to construct the model. Regarding the data, Sumando et al. (2018) used monthly data of government expenditure and developed an individual model for each type of expenditure. Iskandar et al. (2018) utilised only the type of expenditure that is categorised as intermittent expenditure. Unlike Sumando et al. (2018), apart from constructing a model for each type of expenditure, Iskandar et al. (2018) developed a model for the aggregate value of intermittent expenditure. Despite using the monthly data, Iskandar et al. (2018) employed the weekly data of government spending. From the methods point of view, both Sumando et al. (2018) and Iskandar et al. (2018) employed multiple methods to develop a government cash forecasting model and compared the results in order to define the best procedure to construct a reliable cash forecasting model. However, while Sumando et al. (2018) focussed only on the statistical-based methods, Iskandar et al. (2018) explored the use of statistical, machine learning, and hybrid-based models in developing a government cash forecasting. Regardless of what has been done, the previous studies fail to confirm the best procedure to develop a government cash forecasting model. Therefore, the research aim will be addressed by answering the following research questions:

1. What are the most appropriate variables to be included in developing a government cash forecasting model with accuracy that meets an acceptable level of materiality for the cash manager by utilising the historical data of government expenditure?
2. What are the most appropriate techniques to use to develop a government cash forecasting model with accuracy that meets an acceptable level of materiality for the cash manager by utilising historical government expenditure data?

4.4. Summary

The purpose of this chapter was to synthesise the literature reviewed in this study, as a basis for explaining the research aim and the development of the research questions. The importance of having a government cash forecasting model with accuracy that meets an acceptable level of materiality for the cash manager was discussed in Chapter 2. The review of the literature revealed the best practice that a government needs to follow in order to gain an effective GCM. The literature placed the development of a government cash forecasting model as the crucial phase to achieve an effective GCM. Moreover, the review also discovered that the use of the top-down approach to develop a government cash forecasting model is preferable compared to the bottom-up approach. However, there has been little attention given to the top-down approach as an underlying strategy to develop a government cash forecasting model. Upon the basis of this review of the literature several the primary aim and purpose of the research and two related research questions were developed.

The primary research aim elaborates the ultimate objective of this study by questioning the procedure that a government should take in order to develop a government cash forecasting model with accuracy that meets an acceptable level of materiality for the cash manager when the short-term historical data of government expenditure are used. The response to that question can be given by answering the research questions. When it comes to building a forecasting model utilising historical data, the two focal points emerge: the data and the method used. The first research question explores the best data to be used, while the second research question investigates the best method to develop a government cash forecasting model with accuracy that meets an acceptable level of materiality for the cash manager. The following chapter details the data, research design and methods used in the present study.

Chapter 5 Methods

5.1. Introduction

The purpose of this chapter is to detail the research design and methods used to address the research aim and answer the questions developed in Chapter 4. The research aim pursues the exploration of the procedure of developing a government cash forecasting model with accuracy that meets an acceptable level of materiality for the cash manager. The research questions can be answered by identifying and choosing the most appropriate data and methods to be used for constructing the model.

The experimental design of the research is discussed in Section 5.2. The description of data that is used in this research is explained in Section 5.3. The first stage of the research, data pre-processing phase, is discussed in Section 5.4. including the attribute selection process. The second stage of the research, developing forecasting models, is discussed in Section 5.5. where the ARIMA, ANN, and hybrid techniques are described. The last stage of the research, evaluating the performance of the models, is discussed in Section 5.5. A summary for this chapter is provided in Section 5.6.

5.2. Research Design

The research questions of this study posed in Chapter 4 seek the procedure to build a model of cash forecasting for the public sector with accuracy that meets an acceptable level of materiality for the cash manager. In order to achieve this goal, this study emphasises the “top-down” approach proposed by Williams (2009, 2010). The approach suggested the utilisation of historical data gathered from the database managed by the cash manager to develop a government cash forecasting system. However, feeding all the available data into modelling phases does not always deliver a better model (Guyon & Elisseeff, 2003). Moreover, Guyon and Elisseeff (2003) proposed that the variable selection process could improve the performance of the model (Zeng et al., 2015). Following on from this, the first research question asks whether selecting the significant variables to be used to construct a government cash forecasting model could increase the model's performance.

As mentioned in Chapter 2, the methods that are used can be classified into three main approaches: statistical, machine learning, and hybrid technique. Such studies gave evidence of the success of each method in providing a reasonable performance of the forecasting model. Therefore, the second research question asks whether employing a different proposed method to develop a government cash forecasting model could increase the model's performance.

The experiment conducted in this study uses several quantitative methods and settings to ensure the best model is achievable. Overall, it consists of three main stages: (1) Attribute Selection Process; (2) Modelling Process; and (3) Performance Evaluation Process. The purpose of the first step, the attribute selection process, is to identify the significant variables from the available dataset generated from the database. At the conclusion of the first step, two datasets – labelled as Initial Dataset and Selected Attributes – are obtained. The design of this research is summarised in Figure 5.1.

Next, the two datasets are used to develop a government cash forecasting model under the proposed methods in the modelling step. Referring to the previous studies described in Chapter 2, the present study examines the ability of statistical, machine learning, and hybrid methods in establishing a government cash forecasting model. The rationale behind the selection of the proposed methods is to ensure this study utilised methods that represent the leading techniques in solving forecasting issues. By doing it, the robustness of the results will be achievable. Each dataset is fed to the proposed methods and estimated separately to form a model such that for each method, there will be two models produced. One model is constructed with the Initial Dataset while the other is built from the selected attributes.

The performance of the resultant model is then compared with the other to define the best model to forecast government cash needed in the future. This process takes place in the last phase: the performance evaluation process. The dataset that is used to develop the model with the best performance is the answer to the first research question, while its method is the response to the second research question. Overall, the procedure to build the best government cash forecasting model is the answer to the research question of this study.

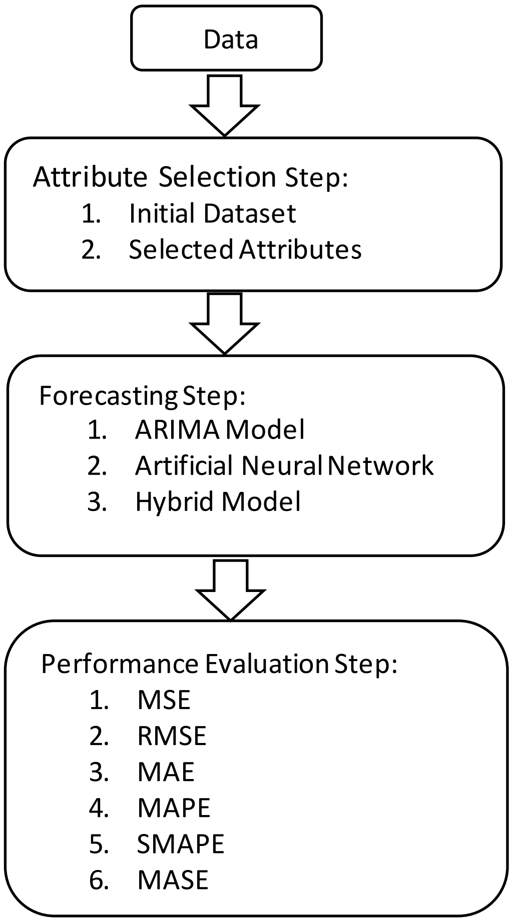


Figure 5.1: Research Design

When dealing with modelling, the famous quotation of *"Essentially, all models are wrong, but some are useful"* from Box and Draper (1987) arose. Reflecting on this, the present study is designed to elaborate the available resources to develop a government cash forecasting model in pursuing the best model with accuracy that meets an acceptable level of materiality for the cash manager. Multiple modelling settings are proposed in order to investigate the most "useful" model of all proposed models, which includes variable selections, forecasting methods, and performance evaluation measurements.

The subject of variable selection to be included in a model has been a consideration of statistical and machine learning modelling studies. A review by Heinze et al. (2018) reported that the variable selection process is seen as a contrast from both modelling approaches. Whilst many statisticians considered that variable selection suffers from an inherent complexity of the model (e.g. issues related to the robustness and the interpretation of the model), it has been seen as “a rule of thumb” for machine learning technique (Heinze et al., 2018). However, there is restriction on both sides of opinion since many studies (e.g. Burnham and Anderson (2002), Royston and Sauerbrei (2008), Steyerberg (2008), Harrell Jr (2015)) suggested that the use of variable selection methods in practice depends on the researchers’ justification and the underlying theories of their studies (Heinze et al., 2018). Hence, to accommodate the contradiction, the present study allows both data from the initial dataset and the selected attributes from the variable selection process to be used in developing a government cash forecasting model. The effect of each data treatment on the accuracy of the model will be definite when the resultant model’s performance is identified. A detailed explanation of the two datasets will be given in Section 5.4.

One of the most frequently used statistical techniques in time series forecasting is Autoregressive Integrated Moving Average (ARIMA). ARIMA represents a generalisation of Autoregressive Moving Average (ARMA) models introduced by Box and Jenkins (1976). Many studies such as Mondal et al. (2014), Ariyo et al. (2014), and Iqbal and Naveed (2016) show that ARIMA has successfully forecast time series data with reasonable accuracy.

Notwithstanding its flexibility in solving the prediction problems, Zhang et al. (1998) pinpointed the assumption of linearity of ARIMA models that limit them in capturing non-linear variations in time series data. Since most of the underlying mechanism of real-world data is nonlinear (Granger &Terasvirta, 1993), utilising ARIMA to develop a government cash forecasting model might be inappropriate. Furthermore, to address this shortcoming, Zhang et al. (1998) proposed ANN models which have proven to have an ability to increase the model’s prediction where the non-linear variations in time series data exist.

As part of a machine learning technique, ANN gained its popularity from its nonlinear forecasting modelling capability (Zhang et al., 2001). Furthermore, the latter study provided evidence of the use of ANN to deal with the nonlinear structure of the data. Studies such as Acuna et al. (2012), Mishra and Dehuri (2014), Venkatesh et al. (2014), and Dandekar and Ranade (2015), revealed the superiority of the ANN model compared to other forecasting tools. Nevertheless, there is no guarantee that ANN always delivers a high level of forecasting accuracy.

Moreover, it is also reasonable to think the data are embedded in both linear or nonlinear patterns. Zhang (2003) stated that it is problematic to decide whether a linear or nonlinear pattern occurs in a time series. With the existence of both patterns in a time series, utilising solely ARIMA or ANN models is not ideal since they are only capable of handling linear and non-linear patterns respectively. Zhang (2003) proposed a combination of ARIMA and ANN models in response to this issue, the so-called hybrid model. The ARIMA and ANN parts of the hybrid model are empowered to analyse the linear and non-linear component of the data. Current studies, such as Wang et al. (2013), Chaâbane (2014), Yu et al. (2014), Moretti et al. (2015), Adhikari (2015), Cadenas et al. (2016), de Oliveira and Ludermir (2016) proved the success of the hybrid model in develop forecasting models. However, Taskaya-Temizel and Casey (2005) argued that the hybrid model does not always deliver better forecasts.

For that reason, the present study utilised ARIMA, ANN, and Hybrid models to identify the best method to describe the structure of the data such that a government cash forecasting model, with accuracy that meets an acceptable level of materiality for the cash manager, is achievable. Both datasets from the attribute selection process, the initial dataset and the selected attributes, are modelled using each proposed method independently. The discussion of the proposed methods in this study will be presented in Section 5.5.

The best forecasting model is decided through the performance evaluation process. When the model is established, each outcome/result is compared following specific model performance evaluation measurements as detailed in Section 5.6. The procedure for developing the best forecasting model is the main aim and outcome of this research.

5.3. Data

The study uses Indonesian central government expenditure data from 2009 to 2015. This is sourced from a database of Indonesian government cash expenditure held by the Indonesian cash manager. Investigating the best way to build a government cash forecasting model, this study focusses on modelling total daily intermittent expenditure of all spending units in Indonesia. Figure 5.2. displays total daily intermittent expenditure of all spending units in Indonesia from 2009 to 2015 summarised for each month of the year.

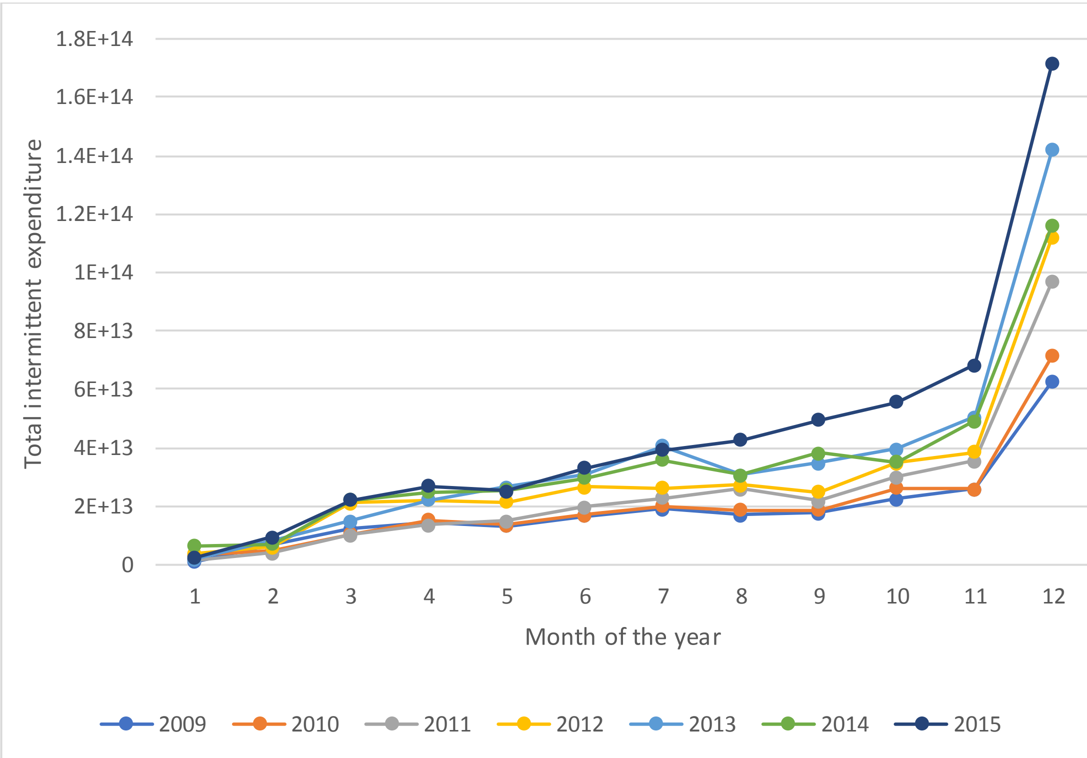


Figure 5.2: Total intermittent expenditure of all spending units in Indonesia from 2009 to 2015

Figure 5.2. shows the patterns of total intermittent expenditure of all spending units in Indonesia. It tends to be very low at the beginning of the budget year, followed by a moderate increase during the second and third quarters of the budget year, and experiencing an extreme intensification at the last month of the budget year.

Although overall patterns display a positive trend during the observation period, the total intermittent expenditure fluctuates in some months. The patterns are almost typical for each budget year.

As described in Chapter 3, total daily intermittent expenditure of all spending units in Indonesia is subject to the constraint of the total daily available fund for intermittent expenditure where maximum daily intermittent expenditure cannot exceed its daily available fund. Moreover, the intermittent expenditure is also influenced by the embedded effect of cash disbursement pattern in Indonesia such as calendar and policy effects. Therefore, this study used data of total daily intermittent expenditure and total daily available fund for the intermittent expenditure of all spending units in Indonesia that are generated from the database. To capture the calendar and policy effects, this study elaborates categorical variables of the day of the week, the week of the month, the month of the year, and policy implementation from its respective date. The categorical variables allow the information to be presented in a number of limited categories (Powers & Xie, 2008). Therefore, the variable for the day of the week consists of five categories representing Monday to Friday, the variable for the week of the month consists of five categories representing Week 1 to Week 5, the variable for the month of the year consists of 12 categories representing January to December, and the variable policy implementation consists of two categories representing before and after new policy implementation.

For experimental purposes, the total daily intermittent expenditure and total daily available fund for intermittent expenditure are transformed to natural logarithms while all categorical variables are coded into dummy variables. Table 5.1. describes the variables that are used in this study.

Table 5.1: List of variables

No.	Variable	Symbol	Type of data	Description
1	Total daily intermittent expenditure	E	continuous data	The total amount of daily intermittent expenditure of all spending units in Indonesia (in natural logarithm form).
2	Total daily available fund for intermittent expenditure	F	continuous data	The total amount of daily remaining budget for intermittent expenditure of all spending units in Indonesia (in natural logarithm form).
3	The day of the week	D	categorical data	Represents the day of the week when the expenditure happened (1=Monday, …., 5=Friday)
4	The week of the month	W	categorical data	Represents the week number of the month when the expenditure happened (1=1st week, …., 5=5th week)
5	The month of the year	M	categorical data	Represents the month of the year when the expenditure happened (1=January, …., 12=December)
6	Policy implementation	$Policy$	categorical data	Represents the new policy implementation (1=Before 2011, 2=2011 and beyond)

The data are split into a training set (2009–2013) and a testing set (2014–2015). The testing set spans a two-year time period to accommodate assessing the effect of budget year shifting in the data. Overall, there are 1,227 observations for training purposes and 488 observations for testing purposes.

5.4. Attribute Selection Process

The ultimate goal of this study is to develop a government cash forecasting model with accuracy that meets an acceptable level of materiality for the cash manager. Previous studies (e.g. Guyon & Elisseeff (2003)) suggested that the performance of the forecasting model can be improved via the variable selection process (Zeng et al., 2015). However, for some forecasting methods, the variable selection process is not always the choice (Heinze et al., 2018). Thus, the present study uses the initial dataset and the selected attributes in order to achieve the best model.

There are six variables that are used in this study. To make sure the best model is achieved, both datasets are employed to construct a government cash forecasting model following the proposed methods independently. The performance of each model built on both datasets is compared to determine whether selecting only the significant variables to develop a government cash forecasting model could increase the performance of the model or not. The following sections are intended to discuss both datasets in more detail.

5.4.1 Initial dataset

Initial dataset is self-explanatory in its definition. It is the original set of data that is not going through any attribute selection process. It includes all variables as described in Section 5.3.

5.4.2 Selected Attributes

In general, there are three attribute selection techniques to consider: filter, wrapper and embedded (De Silva & Leong, 2015). Furthermore, De Silva and Leong (2015) argued that each technique follows a different selection strategy to the observed attributes. The filter technique uses inherent characteristics in the attribute and ranks it based on certain statistical criteria (Eid et al., 2013). On the other hand, wrapper and embedded techniques use the attribute subset selection strategies where the results are presented in a form of a group of attributes. A wrapper model utilises the performance of the learning algorithm as the evaluation criterion, while in embedded techniques the evaluation criterion is built into the learning algorithm itself (De Silva & Leong, 2015).

This study uses attribute selection based on the widely used ranking criterion, the Pearson Correlation Coefficient (PCC), which is part of the filter approach. The main advantage of PCC is its applicability to binary and continuous type attributes which makes the PCC more flexible compared to other techniques (Guyon, 2008). This feature is the rationale for using PCC since this study utilises both binary and continuous data as described in Section 5.3. In addition, utilising the filter approach as an attribute selection tool is faster than other approaches (De Silva & Leong, 2015). For attributes X and Y, the value of PCC r can be calculated using the following formula:

$$r_{XY} = \frac{\sum_{i=1}^{n}(X_i - \bar{X})(Y_i - \bar{Y})}{\sqrt{\sum_{i=1}^{n}(X_i - \bar{X})^2}\sqrt{\sum_{i=1}^{n}(Y_i - \bar{Y})^2}} \qquad 5.1$$

Where $\bar{X}$ and $\bar{Y}$ are the mean of X and Y respectively. The value of PCC is exactly zero when X and Y are independent or uncorrelated, while the values between -1 and +1 represent the level and direction of the correlation. The negative sign represents a negative correlation which means as the variable increases, the independent variable decreases. On the contrary, the positive sign implies the opposite, where an increase in the variable increases will increase the independent variable.

Notwithstanding its simplicity, the shortfall of PCC is that it only measures the linear correlation between variables (Lutu, 2018). When the value of the PCC is exactly or near zero, it does not necessarily indicate that the variable is irrelevant. There is a possibility of a non-linear dependency existing between the observed variables, which is not captured by the PCC (Guyon, 2008). Moreover, Guyon (2008) suggested that transforming the attributes into the natural logarithm form addresses the issue. Referring to the data used by this study as detailed in section 5.3., which is presented in its natural logarithm form, the linearity of PCC is not a concern.

One of the methods to implement the PCC as an attribute selection tool is the Correlation Attribute Evaluation (CAE). The CAE method computes the value of Pearson's correlation between all inputs used in this study to the target. The advantage of using CAE is its ability to calculate the correlation between categorical/nominal and continuous/numeric data. It is done by converting the categorical attributes into its binary indicators before giving the result of overall correlation as a weighted average value (Hall et al., 2009; Frank et al., 2016).

From the perspective of forecasting modelling, an attribute can be defined as a relevant predictor when the attribute is highly correlated to the predictand but at the same time is uncorrelated to other predictors (Hall, 1999). Cohen (1988) argued that the degree of correlation amongst attributes is subject to experimental settings. In the case where the data are collected from "high precision instruments" such as in physical sciences, a coefficient correlation of 0.9 is considered small. However, for social sciences research which requires more complicated experimental setup and measuring instruments, a range value of coefficient correlation is considered. A coefficient correlation of 0.1 – 0.3 indicates a small correlation, 0.3 – 0.5 shows a medium correlation, and 0.5 – 1.00 demonstrates a high correlation (Cohen, 1988). Nevertheless, such classifications are not a rule of thumb for determining whether a coefficient correlation is considered as a high correlation or not since data collection process and experiment setting for each research is incomparable. Therefore, the decision on the attributes that are considered as the relevant predictors is on researcher judgment. Once the attributes are ranked, an appropriate number of attributes are used for the prediction task (De Silva & Leong, 2015).

5.5. Modelling Process

This study employs several different methods and techniques to analyse the data and to develop a model for government cash forecasting. The following section is intended to describe those methods.

5.5.1 Autoregressive Integrated Moving Average

The ARIMA model is a pure time series method in which past values of a variable and a 'white noise' error term is used to forecast future values of the same variable (Gujarati & Porter, 2009). It is developed based on the Box–Jenkins methodology proposed by Box and Jenkins (1976). The most prominent feature of the ARIMA method on developing a forecasting model is the ability to predict the upcoming values of the time series data from their past values without any prerequisite information regarding its underlying theory (Gujarati & Porter, 2009). Such a characteristic gives flexibility to ARIMA in solving challenging forecasting tasks where the relationship amongst the available variables is unclear.

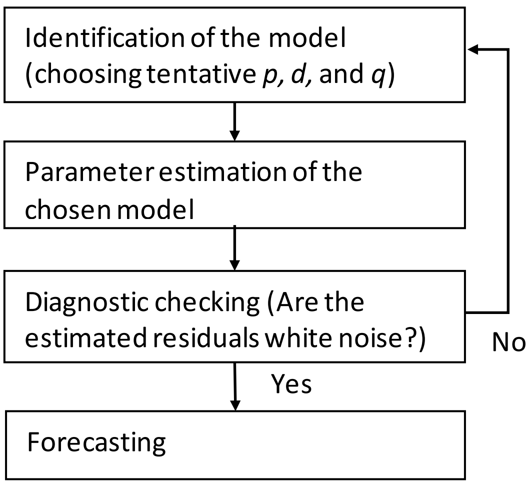
As a statistical method, ARIMA requires the time series data that are used to be stationary. Stationary is a statistical characteristic such that the mean and the autocorrelation structure of a time series data are constant over the observation period (Khashei et al., 2009). Nonstationary time series data will lead to spurious regression problems where the ARIMA model is only capable of estimating phenomena for a specific episode (Gujarati & Porter, 2009). This can be fixed by applying differencing and power transformation to the data to remove the trend and to stabilise the variance (Khashei et al., 2009). The assumption of stationary is necessary to ensure the resultant model has a generalisation ability over the observation periods which is essential for the purpose of forecasting (Gujarati & Porter, 2009).

ARIMA is a modification of an ARMA model (autoregressive moving average) (Mondal et al., 2014). ARMA (p, q) has the form:

$$Y_t = \theta + \sum_{i=1}^{p} \alpha_i Y_{t-i} + \varepsilon_t + \sum_{j=1}^{q} \beta_j \varepsilon_{t-j} \tag{5.2}$$

Y_t is the variable to be predicted based on its preceding p values with constant linear coefficients α and current ε_t and previous errors for q periods with constant linear coefficients β. To ensure a robust time series analysis, the underlying data need to be stationary. If the data are not stationary, they have to be differentiated d times to make them stationary. A difference stationary series is said to be integrated and is denoted as $I(d)$ where d is the order of integration. The order of integration is the number of unit roots contained in the series, or the number of differencing operations it takes to make the series stationary. Therefore, a stationary time series data has no unit root or $I(0)$. When a time series is $I(d)$ and applied to Equation 5.2, the formal ARIMA model is ARIMA(p, d, q), where p, d, and q are the number of autoregressive terms, the order of integration, and the moving average terms respectively and are all nonnegative integers. Consequently, an ARIMA model can be determined once values of p, d, and q are known. For instance, with stationary time series data (d = 0), an ARIMA(p, 0, q) is equal to an ARMA model of ARMA(p, q). Similarly, with the moving average term equal to zero (q = 0), ARIMA(p, 0, 0) is a pure autoregressive process, AR(p), and an ARIMA model are a pure moving average process, MA(q), when the autoregressive term equal to zero (p = 0).

According to Gujarati and Porter (2009), the Box–Jenkins methodology contains four steps: model identification, parameter estimation, diagnostic checking, and forecasting, as shown in Figure 5.3. The basic idea of model identification is that if a time series is generated from an ARIMA process, it should have some theoretical autocorrelation properties. By matching the empirical autocorrelation patterns with the theoretical ones, it is often possible to identify one or several potential models for the given time series (Khashei et al., 2009). Therefore, for ARIMA(p, d, q), Box and Jenkins (1976) proposed the correlograms of autocorrelation functions (ACF) and partial autocorrelation functions (PACF) of the data as the preliminary tools to determine the most appropriate estimates of p and q, while the value of d was decided by implementing a unit root test.



Source: Gujarati and Porter (2009)

Figure 5.3: The Box-Jenkins methodology

The Augmented Dickey-Fuller (ADF) test is the most prominent unit root testing instrument. It is developed from the standard Dickey-Fuller test proposed by Dickey and Fuller (1979). The simple Dickey-Fuller test can be written in mathematical form as follows:

$$\Delta y_t \ = \alpha y_{t-1} + x_t{}'\delta + \varepsilon_t \qquad 5.3$$

where $\alpha = p - 1$. The null and alternative hypotheses may be written as:

H_0: $\alpha =\ 0$, (i.e., there is a unit root or the time series is nonstationary, or it has a stochastic trend). 5.4

H_1: $\alpha < 0$, (i.e., the time series is stationary, possibly around a deterministic trend) 5.5

and evaluated using the conventional t-ratio for α:

$$t_\alpha = \hat{\alpha}/\left(se(\hat{\alpha})\right) \tag{5.6}$$

where $\hat{\alpha}$ is the estimate of α, and $se(\hat{\alpha})$ is the coefficient standard error.

As an improvement of the simple Dickey-Fuller test, ADF overcomes the weakness of the earlier version that is only valid for the time series data with the number of autoregressive terms p equal to one or $AR(1)$. ADF modifies the simple Dickey-Fuller test formula by assuming the time series data y follows an $AR(p)$ process and adding p lagged difference terms of the dependent variable y. Therefore, Equation 5.3 can be represented as follows:

$$\Delta y_t = \alpha y_{t-1} + x_t{}'\delta + \beta_1 \Delta y_{t-1} + \beta_2 \Delta y_{t-2} + \cdots + \beta_p \Delta y_{t-p} + \varepsilon_t \tag{5.7}$$

The ADF test statistic is calculated following Equation 5.6 and evaluated based on the null and alternative hypotheses presented in Equation 5.4 and 5.5, respectively.

The ACF τ_k and PACF ϕ_k value of a time series data at Y at lag k is defined as:

$$\tau_k = \frac{\sum_{t=k+1}^{T} (Y_t - \bar{Y})\,(Y_{t-k} - \bar{Y})}{\sum_{t=1}^{T} (Y_t - \bar{Y})^2} \tag{5.8}$$

$$\phi_k = \begin{cases} \tau_k & , for\ k = 1 \\ \dfrac{\tau_k - \sum_{j=1}^{k-1} \phi_{k-1,j}\, \tau_{k-j}}{1 - \sum_{j=1}^{k-1} \phi_{k-1,j}\, \tau_{k-j}} & , for\ k > 1 \end{cases} \tag{5.9}$$

Where $\bar{Y}$ is the mean of Y and $\phi_{k,j} = \phi_{k-1,j} - \phi_k \phi_{k-1,k-j}$.

The ACF value represents the correlation coefficient of a time series data for each period. A nonzero value means the series is first order serially correlated. When the value is dying off, imitating geometric patterns along with the increasing lag, it is a sign that the series obeys a low-order autoregressive (AR) process. On the contrary, if the value drops to zero after a small number of lags, it is a sign that the series obeys a low-order moving-average (MA) process.

The PACF measures the correlation of values that are periods apart after removing the correlation from the intervening lags. If the pattern of autocorrelation is one that can be captured by an autoregression of order less than k, then the partial autocorrelation at lag k will be close to zero. The PAC of a pure autoregressive process of order p, $AR(p)$, cuts off at lag k, while the PAC of a pure (MA) process asymptotes gradually to zero.

Once the proposed ARIMA configurations are identified, the next step is to estimate the models and select the best model based on their information criteria values. Akaike Information Criterion (AIC), Bayesian Information Criterion (BIC), and Hannan-Quinn Information Criterion (HQC) proposed by Akaike (1987), Schwarz (1978), and Hannan and Quinn (1979) respectively are amongst the list of information criteria that are commonly used for ARIMA Model selection. The smallest value of the information criteria is the best model to represent the real data. AIC, BIC, and HQC are calculated following the formulae as shown below (IHS Global, 2016):

$$AIC = -2(l/T) + 2(k/T) \tag{5.10}$$

$$BIC = -2(l/T) + k\log(T)/T \tag{5.11}$$

$$HQC = -2(l/T) + 2k\log\big(\log(T)\big)/T \tag{5.12}$$

Let l be the natural logarithm form of the value of the likelihood function with the k parameters estimated using T observations. The various information criteria are all based on -2 times the average log likelihood function, adjusted by a penalty function. Since the underlaying theory behind each information criterion is different, it is reasonable to expect disagreement amongst the criteria regarding the best model to select. On such an occasion, the best model is chosen based on the one with the best forecasting ability.

To make sure the selected model is fit for the data, a white noise test on residuals is employed in the diagnostic checking phase. It is done by obtaining the correlogram of the ACF and PACF of its residuals. The selected model is considered as the best fit for data when the residuals estimated are purely random which is shown by its ACF and PACF values falling within the significance level boundaries. When the residuals are not white noise, the process repeats from the first step iteratively.

Once the selected model is confirmed, the last phase is deployed. The forecasting step produces the predicted future values of the time series data. The forecasting performance of the model is measured using the proposed performance evaluation metrics described in Section 5.6. The resultant performance is then compared to other forecasting performances from the other proposed methods for the best model identification purposes.

In the case of multivariate analysis, the ARIMA model as described in Equation 5.2 can be modified into an ARIMA with Exogenous Variable (ARIMAX) which is written in mathematical form as follows (Neshat et al., 2018):

$$Y_t = \theta + \sum_{i=1}^{p} \alpha_i Y_{t-i} + \varepsilon_t + \sum_{j=1}^{q} \beta_j \varepsilon_{t-j} + \sum_{k=1}^{n} \gamma_k X_k \tag{5.13}$$

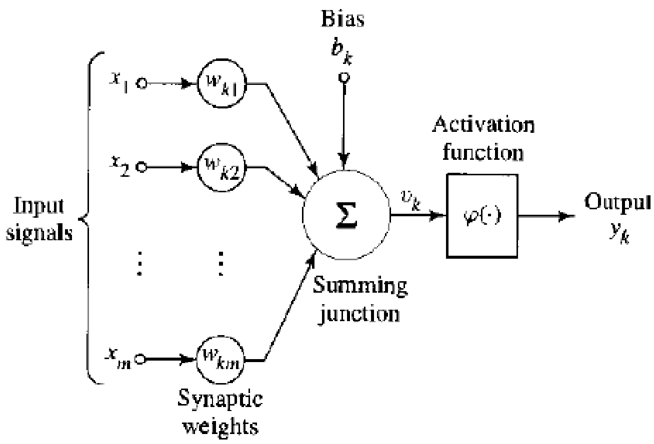
Where X and γ denote the exogenous variable and its constant linear coefficients, respectively, while n is the number of exogenous variables.

Furthermore, Wei (2006) argued that many businesses and economic time series data are embodied with repeated phenomena over a period of time. This regular period of time is termed as the seasonal period. Some research has suggested the use of Seasonal ARIMA (SARIMA) to deal with such phenomena. However, the Pennsylvania State University (2018) stated that SARIMA only works when the seasonal period is in a regular pattern.

Research by Cools et al. (2009) provided evidence for the case of daily forecasting, that the ARIMAX technique performs better than the SARIMA model in capturing seasonality of the data. Anggraeni et al. (2015) considered the ARIMAX model as an alteration of ARIMA to handle seasonality. Such modification is done by incorporating the calendar effects into predictor variables. Therefore, this study uses the ARIMAX method to predict the future value of total daily intermittent expenditure with its total daily available fund, the calendar effects, and policy implementation described in Section 5.3. as the exogenous variables.

5.5.2 Artificial Neural Network

ANN is a computational model inspired by human brain function. It was developed to mimic biological neural systems as an information processing system. The major advantage of ANN compared to other methods is the ability to learn from experience (Zhang et al., 1998). According to Yildiz and Yezegel (2010), the main elements of ANN are neurons, connections, and the learning algorithm. Similar to the brain, ANN has a neuron as an information-processing unit. As shown in Figure 5.4., a neuron consists of a set of connections called synapses. Each synapse has a weight value that represents the strength of the respective signal. In a neuron k, an input data x_j at synapse j is multiplied by the synaptic weight w_{kj}. This product of all synapses is then summed up into a single value called linear combiner v_k. It includes a bias b_k which has the effect of increasing or decreasing the net input of the activation function $\varphi(.)$ (Haykin, 1999). The output of the neuron y_k can be used by other neurons (Butler, 2006).



Source: (Haykin, 1999)

Figure 5.4: A nonlinear model of a neuron

In mathematical terms, a neuron k may be written as follows:

$$y_k = \varphi(v_k) \tag{5.14}$$

Where

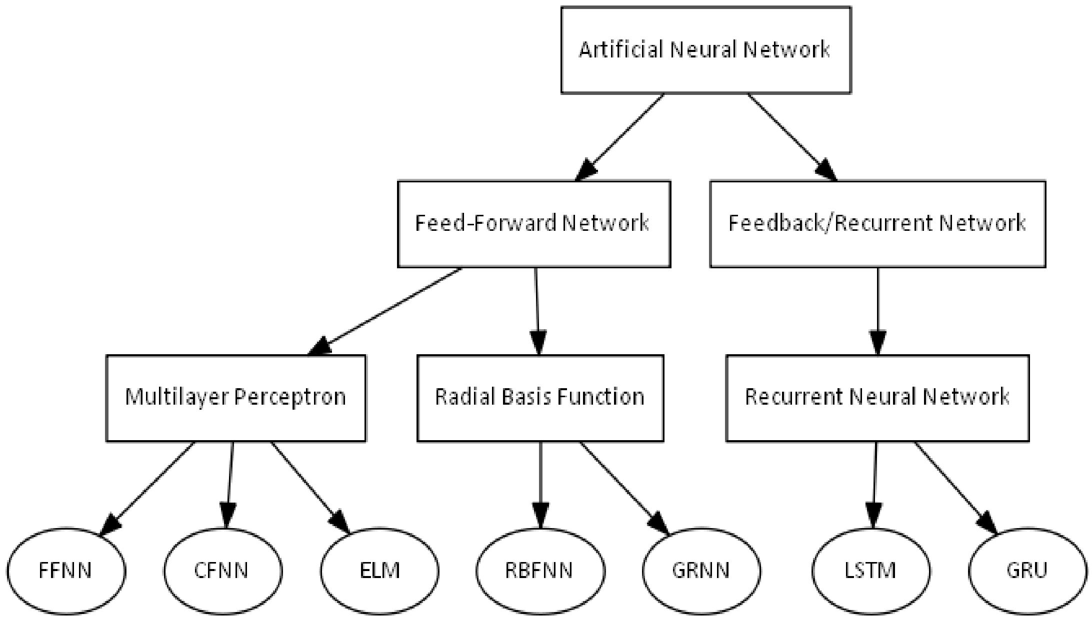
$$v_k = \sum_{j=1}^{m} w_{kj} x_j + b_k$$

A neuron's activation function may follow a threshold function, a piecewise-linear function, or a sigmoid function (Haykin, 1999). The Sigmoid function is the most frequently used in ANN (Butler, 2006; Yildiz & Yezegel, 2010).

According to Haykin (1999), ANN simulates the human brain in two ways: (1) acquiring knowledge through a learning process from its environment and (2) storing it into an interneuron connection strength known as synaptic weight. The learning process is established through a learning algorithm, which is a procedure to adjust the synaptic weight in order to achieve the expected goals. A repetitive process is used until the desired outputs are produced. The network is considered trained when this process is concluded (Yildiz & Yezegel, 2010).

From the learning algorithm perspective, ANN can be categorised into three primary learning paradigms, namely supervised, unsupervised, and hybrid learning. A supervised learning paradigm gives the network output target value such that the proposed weights are dedicated to providing an output as close as possible to the target set. On the other hand, in unsupervised learning, there is no target output associated with each input during the network training. Lastly, a hybrid learning regime combines both supervised and unsupervised learning. Via hybrid learning, some of the weights are defined through supervised learning while the rest are following an unsupervised learning framework of Jain et al. (1996). Due to the nature of the data described in Section 5.3., the present study employs ANN based on a supervised learning paradigm.

To address the objective, this study has developed a number of forecasting models following the proposed ANN techniques namely Feed-forward Neural Networks (FFNN), Cascade-forward Neural Networks (CFNN), Radial Basis Function Neural Network (RBFNN), Generalised Regression Neural Network (GRNN), Extreme Learning Machine (ELM), Long Short-Term Memory (LSTM), and Gated Recurrent Unit (GRU). The proposed ANN techniques are intended to cover most of the ANN classification such that the best procedure in developing a government cash forecasting model is definite. Based on architectural structures, ANN can be classified into (1) feed-forward networks, such as multilayer perceptron and radial basis function based neural networks, and (2) feedback/recurrent networks such as a recurrent neural network (Jain et al., 1996). Feed-forward networks may be distinguished from recurrent/feedback networks based on the direction of their connection. Notwithstanding the fact that the feed-forward neural networks (FFNN) are the most commonly used for forecasting tasks (Zhang et al., 1998, Basu et al., 2010), this study utilises some ANN techniques that cover most of the classification above to investigate the best forecasting model. Figure 5.5. states the proposed ANN techniques position based on its architectural structures.



Adapted from Jain et al. (1996)

Figure 5.5: The ANN classification based on architectural structure

ANN can also be classified based on the depth of the network, which is shallow learning and deep learning. According to Pascanu et al. (2013), the depth of ANN is determined from the number of nonlinear layers which are passed from the input information to output. By this definition, a shallow learning network can be seen as an ANN with less hidden layers than the one in a deep learning network. A recent study by Lv et al. (2015) suggested that the deep architecture of ANN is more powerful models than shallow networks for prediction purposes. In this sense, the class of feedback/recurrent networks is deeper than the feed-forward networks-based ANN structure due to its ability in processing both sequential and parallel information in a more natural and efficient way (Schmidhuber, 2015). The loops that are created during the feedback/recurrent networks process can be expressed as a composition of multiple nonlinear layers when unfolded in time (Pascanu et al., 2013). Latest deep architectures use several modules that are trained separately and stacked together so that the output of the first one is the input of the next one (Gamboa, 2017). Figure 5.6. outlines the classification of the proposed ANN techniques based on the depth of the ANN architectures. The following sections detail the ANN techniques used in this study.

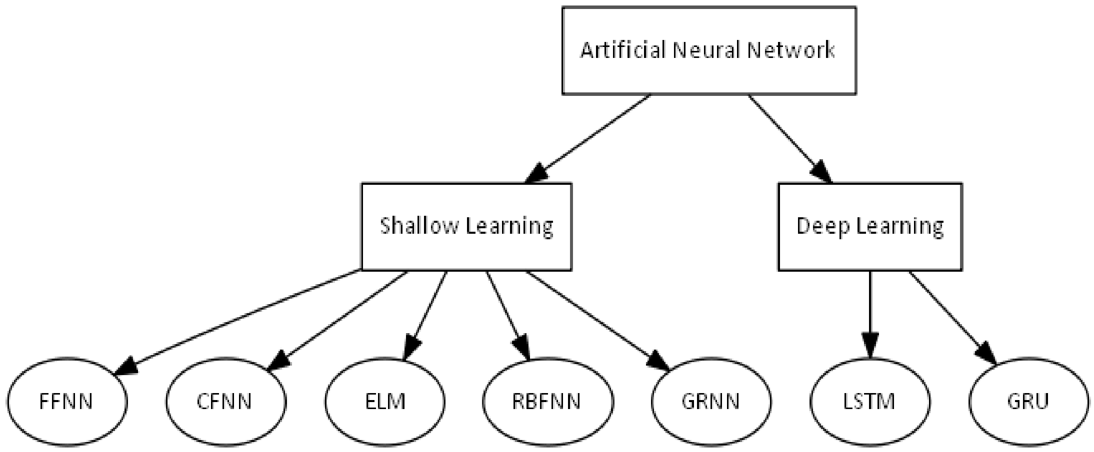
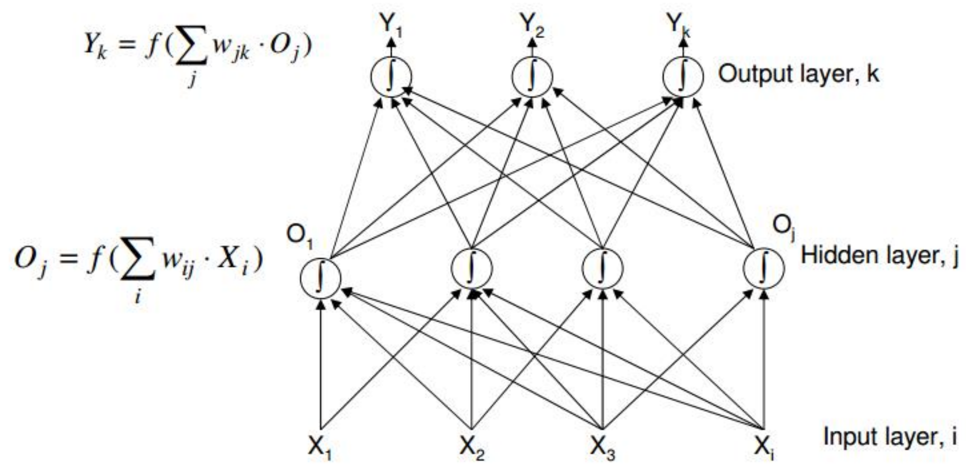


Figure 5.6: The ANN classification based on the depth of the ANN structure

5.5.2.1. Feed-forward Neural Network

FFNN, one of the MLP architectures, is the most common feed-forward network architecture. It consists of input, output, and at least one hidden layer (Sajjadi et al., 2016). The presence of the hidden layer enables FFNN to continuously learn from previous information via a high degree of connectivity, determined by the synapses of the network (Haykin, 2009). The information propagates forward from the input nodes to the output nodes, sequentially without any loops (Ceniceros et al., 2015). Behbahani et al. (2018) listed the major characteristics for MLP/FFNN architecture as follows: the number of neurons per hidden layer is independent of the number of input or output neurons; output and input layers are connected via the hidden layer; neurons in a layer are not connected to each other, while they are fully connected between layers; and there is no requirement for the number of output and input units to be equal. A typical multilayer FFNN is shown in Figure 5.7.

As a network, FFNN relies on how the neurons are connected. Figure 5.7. illustrates a simple structure of a fully connected FFNN that includes an input layer, a hidden layer, and an output layer. The term “layer” indicates the row of the neurons in the structure of FFNN (Yildiz & Yezegel, 2010). An input layer receives data from outside the neural network and sends it to the hidden layer in the weighted form. Hidden layers are layers between input and output. An FFNN may have more than one hidden layer. In those layers, the weighted signal received from the input layer is processed following its nonlinear transfer function as an activation function and then transmitted to the output layer. The output layer obtains information from each neuron in the hidden layer in the form weighted sum which is then processed by the output layer’s activation function. An output layer sends calculated results outside the ANN. The learning process takes place in the output layer. Results obtained from this process are the output values for the problem handled by the FFNN. The number of neurons in the input layer and the output layer is equal to the number of input data and data received from the hidden layer respectively (Staub et al., 2015, Warsito et al., 2018).



Source: Bishop (1995)

Figure 5.7: A simple FFNN

Denote the output of a hidden neuron of an FFNN O_j and the output of an FFNN Y_k as follows, respectively,

$$O_j = f\left(\sum_{i=0}^{m} w_{ij}^{(h)} X_i\right) \tag{5.15}$$

$$Y_k = f\left(\sum_{j=0}^{q} w_{jk}^{(d)}\, O_j\right) \qquad 5.16$$

where $w_{ij}^{(h)}$ is the weight of the j^{th} node in the hidden layer h, connected to the i^{th} node from the previous layer. X_i is the input signal, $w_{jk}^{(d)}$ is the weight of k^{th} node in the output layer d connected to the j^{th} node from the previous layer and $f(.)$ is a nonlinear activation function in each neuron. A general form of FFNN is specified from Equations 5.15 and 5.16 as follows (Bishop, 1995):

$$Y_k = f\left(\sum_{j=1}^{q} w_{kj}^{(d)}\, f\left(\sum_{i=1}^{m} w_{ji}^{(h)}\, X_i\right)\right) \qquad 5.17$$

In the presence of bias on each layer, Equation 5.17 can be written as follows:

$$Y_k = f\left(\sum_{j=1}^{q} w_{kj}^{(d)}\, f\left(\sum_{i=1}^{m} w_{ji}^{(h)}\, X_i + b_k^{(h)}\right) + b_k^{(d)}\right) \qquad 5.18$$

Where $b_k^{(d)}$ and b_k^{h} are the weight of the bias of the output and hidden layer, respectively.

The most popular learning algorithm in a feed-forward neural network is the back-propagation (BP) algorithm. It uses the back propagation of the error gradient in order to achieve the best fit of FFNN architecture (Yan & Zou, 2013). According to Haykin (2009), the BP learning algorithm contains a forward pass and a backward pass. First, a forward pass processes the information from the input node layer by layer over the network in a forward direction and the network produces an output as a response to this process. Then, the BP algorithm redistributes the respective error of the output back through the network called the backward pass process. It adjusts the synaptic weights simultaneously until the minimum error is achieved, through several iterations. Despite its simplicity to determine the weights in the network, the traditional BP algorithm suffers from slow computation time (Yan & Zou, 2013) and requires a great deal of training data (Panchal et al., 2011). The present study will use the Levenberg-Marquardt (LM) algorithm, proposed by Levenberg (1944) and Marquardt (1963), which has proved to be faster than the traditional BP algorithm (Ebtehaj & Bonakdari, 2016; Dudek, 2016; Ayala & Coelho; 2016, Wang et al., 2015).

The Levenberg-Marquardt method combines Newton's method and the gradient descent-based algorithm. While relatively fast in learning the information from the data, the Newton's method-based algorithm sometimes fails to provide an optimum result. On the other hand, the gradient descent method guarantees a proper selection of the parameters but converges slowly (Haykin, 2009).

Let the function $F(w)$ denote the input-output mapping of an FFNN trained with the BP algorithm that follows:

$$F(w) = e^T e \quad 5.19$$

where $w = [w_1, w_2, \cdots, w_N]$ are all weights of the network, e is the error vector comprising the error for all the training examples. The learning algorithm aims to optimise the weight correction Δw. Newton's method calculates the optimum value of the adjustment correction Δw as

$$\Delta w = \mathrm{H}^{-1}\mathrm{g} \quad 5.20$$

where g is the local gradient vector, defined by

$$\mathrm{g} = J^T e \quad 5.21$$

H is its Hessian matrix, defined by

$$\mathrm{H} = J^T J \quad 5.22$$

where J is the Jacobian matrix that contains the first derivatives of the network errors with respect to the weights and biases. The Jacobian matrix can be computed through a standard backpropagation technique that is much less complex than computing the Hessian matrix (Beale et al., 1992).

In general, the gradient descent-based method estimates the optimum value of the adjustment correction Δw by

$$\Delta w = \alpha \mathrm{g} \quad 5.23$$

Where α determines the learning rate parameter of the gradient descent-based algorithm and a direction for weight change that reduces the value of error.

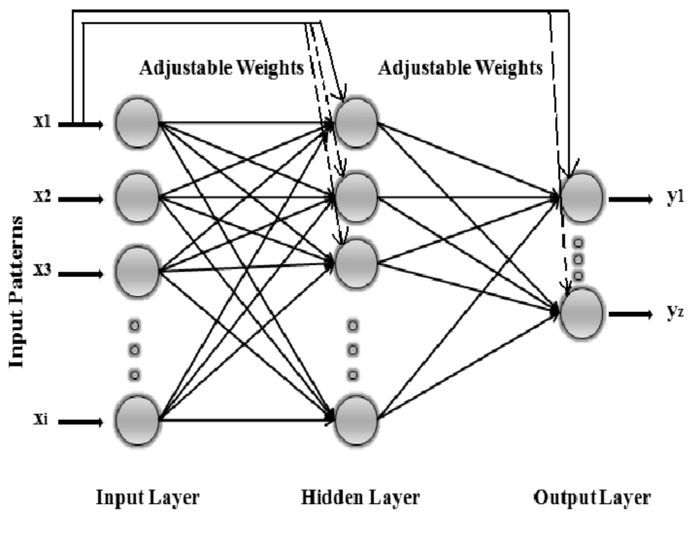
The Levenberg-Marquardt algorithm acts like the gradient-descent and the Gauss-Newton method alternately depends on how close the parameters are to their optimal value (Gavin, 2017). In the LM algorithm, the weight correction Δw is calculated as follows:

$$\Delta w = [\mathrm{H} + \lambda\mathrm{I}]^{-1}\mathrm{g} \qquad 5.24$$

Where I is the identity matrix of the same dimensions as H, and λ is a controlling parameter, also called learning rate, that forces the sum matrix $[\mathrm{H} + \lambda\mathrm{I}]$ to be positive definite which is required for the computation. When the value of λ is close to zero, the LM algorithm follows Newton's method while on the contrary, large values of λ result in a gradient descent update.

5.5.2.2. Cascade-forward Neural Network

As part of the MLP class, CFNN shares the same architectural structure with FFNN in terms of feed-forward connections within the network. Warsito et al. (2018) consider CFNN as the special class of the FFNN model where there is an additional connection from the input and every previous layer to the next layers. Therefore, a CFNN with three layers has the same connections with FFNN with an extra connection from the first layer to the last layer, for instance. The extra connection in its network made CFNN a more complex form of models compared to FFNN (Biswas et al., 2014). However, Beale et al. (1992) argue that the additional connections are expected to enhance the speed of the network to learn the relationship amongst data. CFNN has the ability to capture the nonlinear relationship between input and output while keeping the linear relationship between them at the same time (Warsito et al., 2018). A simple CFNN with one hidden layer is presented in Figure 5.8.



Source: Biswas et al. (2014)

Figure 5.8: A simple CFNN

Since the CFNN is the FFNN with extra weight connections, the output of the k-th neuron, y_k in the output layer of CFNN can be derived from Equation 5.18 as follows (Al-Batah et al., 2015):

$$y_k = f\left(\sum_{j=1}^{J} w_{jk}\ f\left(\sum_{i=1}^{I} w_{ij}x_i + b_j\right) + \left(\sum_{i=1}^{I} w_{ik}x_i + b_j\right) + b_k\right) \qquad 5.25$$

Where k, j,and i are the output node, the hidden node, and the input node, respectively. w and b are the weight and bias, respectively. For training purposes, this study uses the LM method as the learning algorithm to the CFNN.

5.5.2.3. Radial Basis Function Neural Network

Similarly to MLP, RBFNN is composed of the input, hidden and output layers (Ayala & Coelho, 2016). However, in RBFNN, a single hidden layer with neurons of the radial basis activation functions and an output layer that is a weighted sum of the hidden neurons are used (Reiner, 2015). Despite the fact that RBFNN requires more neurons, its training time is generally faster than standard feedforward networks (Beale et al., 1992). An example of RBFNN architecture can be seen in Figure 5.9.

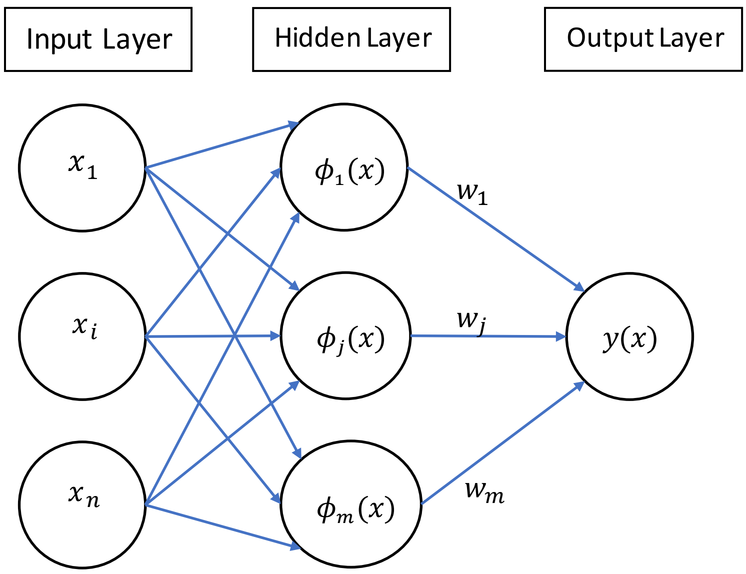


Figure 5.9: A simple RBFNN

An output of RBFNN $y(x)$ can be calculated following the mathematical equation as follows:

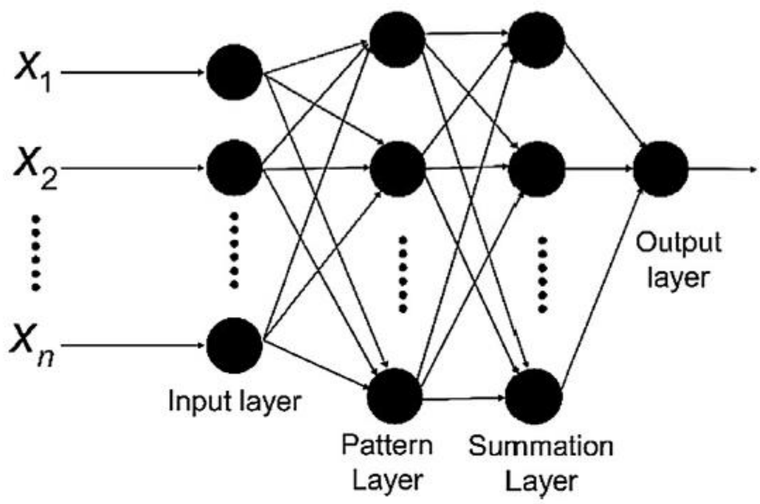
$$y(x) = \sum_{j=1}^{m} w_j \phi_j(x) \quad 5.26$$

Where w_{kj} is output weight, m is the number of hidden layer neurons, and ϕ_j is the radial basis activation function. Haykin (2009) listed the potential activation function for the RBF network based on a theorem proposed by Micchelli (1984). The list includes Multiquadric, Inverse Multiquadric, and Gaussian functions (Haykin, 2009). Nonetheless, the most commonly used radial basis activation function is the Gaussian basis function which is calculated using Equation 5.27:

$$\phi_j(x) = exp\left(-\frac{\|x-\mu_j\|^2}{2\sigma_j^2}\right) \quad 5.27$$

where σ_j represents the spread of the radial basis function and the norm $\|\cdots\|$ denotes the Euclidean distance between the set of inputs, x and the centre of RBF, μ_j.

5.5.2.4. Generalised Regression Neural Network



Source: Feng et al. (2017)

Figure 5.10: Structure of GRNN

First proposed by Specht (1991), GRNN is grouped as a radial basis function network class (Chin et al., 2017; Feng et al., 2017). As shown in Figure 5.10, it comprises of input layer, pattern layer, summation layer and output layer (Yip et al., 2014; Panda et al., 2015). Whilst its input layer and output layer are the same, GRNN differs from RRBFNN in its intermediate layers (Modaresi et al., 2018). In the pattern layer, a neuron is assigned to each input vector following radial basis activation function, which is the Gaussian function as described in Equation 5.27. Therefore, the number of neurons assigns for GRNN is equal to the number of instances in the input layer. Next, its outputs are multiplied with specified connection weights and summed in the summation layer. Once the result is definite, the outputs of the summation layer are fed into the output layer (Panda et al., 2015). The ability to develop a model with a relatively faster learning time (Panda et al., 2015) and an acceptable performance are some of the main advantages of GRNN (Al-Mahasneh et al., 2018).

The estimated GRNN can be calculated following the formula presented in Equation 5.28 (Al-Mahasneh et al., 2018):

$$y(x) = \frac{\sum_{i=1}^{n} y_i \exp(-D_i/2\sigma^2)}{\sum_{i=1}^{n} \exp(-D_i/2\sigma^2)} \qquad 5.28$$

where

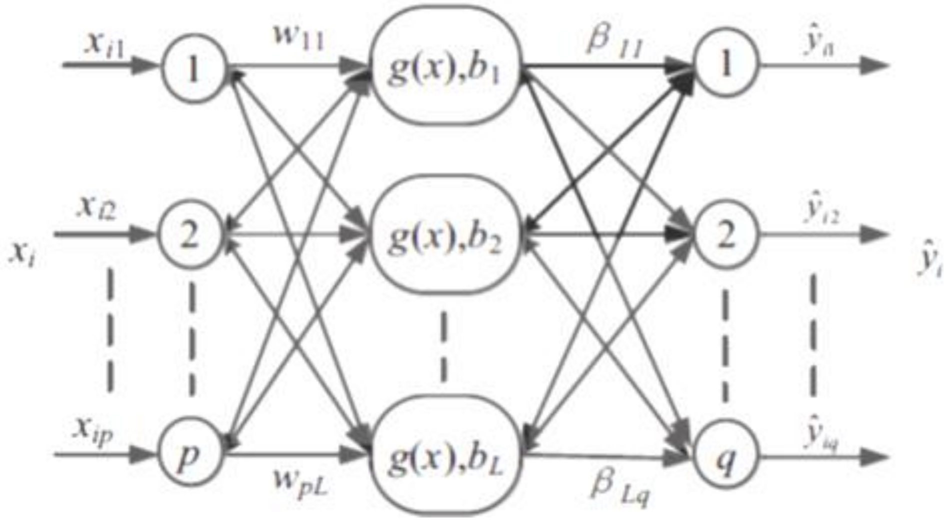
$$D_i = (X - X_i)^T . (X - X_i) \qquad 5.29$$

σ and n are the smoothing parameter and the number of samples, respectively. The estimated value $y(x)$ is calculated by averaging the weighted value of the sample observations y_i, which is the exponential of the squared Euclidian distance between samples X and X_i (Panda et al., 2015).

5.5.2.5. Extreme Learning Machine

Notwithstanding its popularity, FFNN has a major concern regarding its learning speed and optimisation problem (Chen et al., 2012, Huang et al., 2015). According to Huang et al. (2015), such conditions are caused by the use of the gradient descent-based learning algorithms, as the most common learning methods of feed-forward neural networks. Despite all alternative techniques (e.g. second order optimisation methods, subset selection methods, or global optimisation methods) developed to increase the efficiency in training FFNN, there is no guarantee for an optimal global solution (Huang et al., 2015).

An extreme learning machine (ELM) was proposed by Huang et al. (2004) to overcome the feed-forward neural network's issue as mentioned above (Chen et al., 2012). ELM is a single hidden layer feed-forward network. Unlike the use of the BP and the LM learning algorithms as in the case of the other MLP techniques, where the hidden and output parameters are determined by following an optimisation process such as the gradient descent-based optimisation, ELM generates Its hidden nodes (input weights and biases) randomly and then fixed without an iterative process (Chen et al., 2017). The only parameters which need to be learned are the output weights (Huang et al., 2006). Compared to traditional MLP learning methods, ELM has been theoretically proven to be able to provide a universal approximation, good generalisation capabilities and extremely fast speed (Lolli et al., 2017). The key was the ability of ELM to handle problematic issues of nonlinear optimisation (e.g. the optimal determination for input weights, hidden layer biases, output weights) in traditional FNN algorithms with a simple least squares problem of deciding the optimal output weights (Chen et al., 2017; Zhang et al., 2017). Furthermore, Chen et al. (2017) argued that a better generalisation performance of ELM comes from an iteration process of the feedforward networks to reach the smallest training error, but also the smallest output weight norm. Figure 5.11 describes the ELM network.



Source: Zhang et al. (2017)

Figure 5.11: Structure of ELM

In a mathematical sense, ELM can be described as follows (Zhang et al., 2017, Huang et al., 2004). For N arbitrary distinct input variable X, where $X = [x_1, x_2, \dots, x_N] \in \mathbb{R}^{N\times p}$ and output variable $Y = [y_1, y_2, \dots, y_N] \in \mathbb{R}^{N\times p}$, standard SLFN with L hidden nodes and activation function $g(x)$ are modelled as

$$\sum_{i=1}^{L} \beta_i g_i(\mathbb{x}_j) = \sum_{i=1}^{L} \beta_i g(\mathbb{w}_i \mathbb{x}_j + b_i) = \hat{y}_j, \qquad j = 1, \dots, N \tag{5.30}$$

where $\mathbb{w}_i$ is the weight vector connecting the ith hidden node and the input nodes, β_i is the weight vector connecting the ith hidden node and the output nodes, and b_i is the bias of the ith hidden node.

The aim of ELM training is to find the optimal value of $\mathbb{w}_i$, b_i, and β_i that satisfies the cost function defined by

$$f(\mathbb{w}_i, b_i, \beta_i) = \sum_{j=1}^{N} \|\hat{y}_j - y_j\|, \qquad i = 1, \dots, L \tag{5.31}$$

The minimum value of cost function can be further expressed as:

$$\min f(\mathbb{w}_i, b_i, \beta_i) = \min \|H(w_1, \cdots, w_L, b_1, \cdots, b_L, x_1, \cdots, x)\beta - Y\| \tag{5.32}$$

Where H is the transformation output matrix of the hidden layer, β is the output weight, and Y is the desired value matrix of the samples. Equation 5.32. can be written compactly as:

$$H\beta = Y \tag{5.33}$$

Where

$$\begin{aligned} H &= H(w_1, \cdots, w_L, b_1, \cdots, b_L, x_1, \cdots, x) \\ &= \begin{bmatrix} h(x_1) \\ \vdots \\ h(x_N) \end{bmatrix} = \begin{bmatrix} g(\mathbb{w}_1 \mathbb{x}_1 + b_1) & \cdots & g(\mathbb{w}_L \mathbb{x}_1 + b_L) \\ \vdots & \cdots & \vdots \\ g(\mathbb{w}_1 \mathbb{x}_N + b_1) & \cdots & g(\mathbb{w}_L \mathbb{x}_N + b_L) \end{bmatrix}_{N \times L} \end{aligned} \tag{5.34}$$

$$\beta = \begin{bmatrix} \beta_1 \\ \vdots \\ \beta_L \end{bmatrix}_{L \times q} \text{ and } T = \begin{bmatrix} \hat{y}_1 \\ \vdots \\ \hat{y}_N \end{bmatrix}_{N \times q} \tag{5.35}$$

The ELM theories claim that the network cost function can be minimised by randomly assigning the hidden nodes' learning parameters w_i and b_i without considering the input data such that Equation 5.34 becomes a linear system. Therefore, the output weights β can be analytically determined by finding a least square solution as follows:

$$\hat{\beta} = H^{\dagger} Y \tag{5.36}$$

where $H^{\dagger}$ is the Moore–Penrose generalised inverse of H calculated as $(H^T H)^{-1} H^T$.

Equation 5.36 shows how a mathematical transformation generates the output weights of ELM. This calculation avoids the parameters of the network to be adjusted iteratively with some appropriate learning parameters such as learning rate and iterations, which cause a long training phase (Ding et al., 2014). ELM algorithm can be summarised as follows (Huang et al., 2004):

Input: a training set $(\mathbb{x}_i, y_i) \in \mathbb{R}^n \times \mathbb{R}^m$ $(i = 1, \ldots, N)$, the activation function $g(x)$, and the hidden layer node number L.

Output : the output weights β.

Step 1. Randomly assign the parameters of hidden nodes $(w_i, b_i), i = 1,2, \ldots, L$

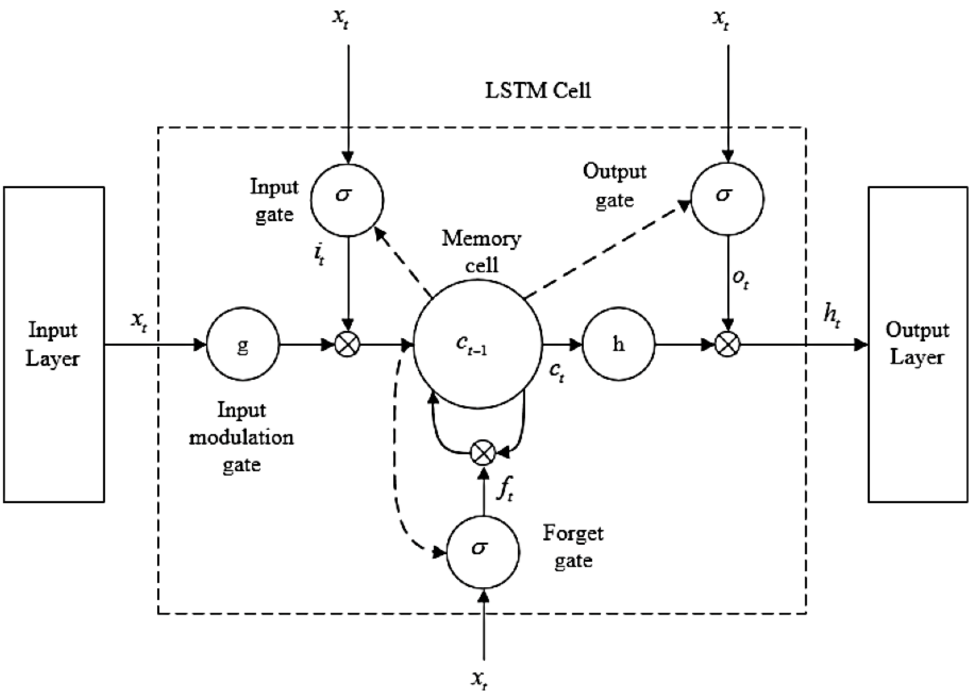
Step 2. Calculate the output matrix of the hidden layer H.

Step 3. Calculate the output weight: $\beta = H^{\dagger}T$.

5.5.2.6. Long Short-Term Memory

The LSTM is part of the recurrent neural networks (RNN) class architecture. A typical feed-forward neural networks architecture contains many connected processing neurons flowing from the input layer to the output layer via dedicated hidden layers. Each neuron is activated in the reception of weighted connections from previously active neurons. The representation of each input event streams from the input layer to the output layer in one direction only. On the other hand, RNN architecture permits a feedback connection to exist which allows the networks to store the recent input information in the form of activations (Schmidhuber, 2015). The capability to learn highly complex vector-to-vector mappings by preserving a vector of activations for each timestep causes the recurrent networks to be considered as deep learning neural networks (Jozefowicz et al., 2015).

However, as mentioned by Hochreiter (1991) and Bengio et al. (1994), the conventional algorithm, such as Back-Propagation Through Time (BPTT) and Real-Time Recurrent Learning (RTRL), leads the RNN to suffer from the exploding and the vanishing gradient problems (Jozefowicz et al., 2015). In addressing the weaknesses, Hochreiter and Schmidhuber (1997) proposed an LSTM that employed an “internal states of special units” called a memory cell which allows enforcing constant error flow in constructing an architecture. The LSTM memory cell allows the network to study when to discard information from the previous hidden states and to update the given hidden states with new information (Graves & Jaitly, 2014). A simple LSTM memory cell is shown in Figure 5.12.



Source: Fu (2016)

Figure 5.12: An illustration of LSTM memory cells

Denote the input time series as $X = (x_1, x_2, \ldots, x_t)$, hidden state of memory cells as $H = (h_1, h_2, \ldots, h_t)$, output time series as $Y = (y_1, y_2, \ldots, y_t)$. LSTM can be written in the mathematical form as follows (Fu, 2016):

$$h_t = H(W_{hx}x_t + W_{hh}h_{t-1} + b_h) \qquad 5.37$$

$$p_t = W_{hy}y_{t-1} + b_y \qquad 5.38$$

Where W is the weight matrices, b is the bias vectors, and p is the predicted values. The hidden state of memory cells, H, is computed in the following formula (Graves & Jaitly, 2014, Graves, 2013, Donahue et al., 2015):

$$i_t = \sigma(W_{xi}x_t + W_{hi}h_{t-1} + W_{ci}c_{t-1} + b_i) \qquad 5.39$$

$$f_t = \sigma\left(W_{xf}x_t + W_{hf}h_{t-1} + W_{cf}c_{t-1} + b_f\right) \qquad 5.40$$

$$o_t = \sigma(W_{xo}x_t + W_{ho}h_{t-1} + W_{co}c_t + b_o) \qquad 5.41$$

$$g_t = tanh(W_{xc}x_t + W_{hc}c_{t-1} + b_c) \qquad 5.42$$

$$c_t = f_t c_{t-1} + i_t g_t \qquad 5.43$$

$$h_t = o_t tanh(c_t) \qquad 5.44$$

Where i_t, f_t, o_t, g_t and c_t are the input gate, forget gate, output gate, input modulation gate, and memory cell unit respectively, all of which are the same size as the hidden vector h_t. σ is the standard logistic sigmoid function where $\sigma(x) = (1 + e^{-x})^{-1}$ and $tanh$ is the hyperbolic tangent function where $tanh(x) = (e^x - e^{-x})/(e^x + e^{-x})$.

The memory cell unit c_t is a sum of two terms: (1) the previous memory cell unit c_{t-1} which is modulated by f_t, and (2) g_t, a function of the current input and previous hidden state, modulated by the input gate i_t.

Because i_t and f_t are sigmoidal, their values lie within the range [0; 1], and i_t and f_t can be thought of as knobs that the LSTM learns to selectively forget its previous memory or consider its current input. Likewise, the output gate f_t learns how much of the memory cell to transfer to the hidden state. These additional cells seem to enable the LSTM to learn complex and long-term temporal dynamics for a wide variety of sequence learning and prediction tasks. Additional depth can be added to LSTMs by stacking them on top of each other. A network with two LSTMs is presented in Figure 5.13. as an illustration.

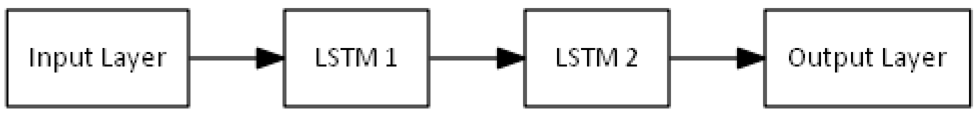
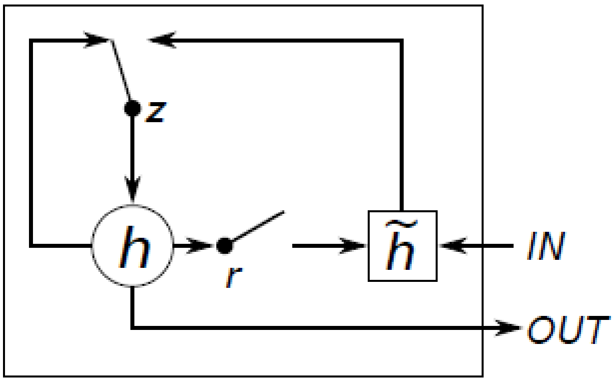


Figure 5.13: An illustration of stacked LSTM

5.5.2.7. Gated Recurrent Unit

Notwithstanding its success in handling the issue described in the previous section, LSTM has received criticism regarding its training time, which is considered slow, (Fu, 2016) and the potential trap of local optima (Jozefowicz et al., 2015). Encountering these challenges, Cho et al. (2014) initiated a novel RNN technique called Gated Recurrent Units (GRU). Similar to LSTM, in a GRU network the information flows are modulated inside the gating units. Nevertheless, the absence of separate memory cells differentiate it from LSTM (Chung et al., 2014).

There are two gates in a GRU: reset gate and update gate. The reset gate allows the network to ignore the irrelevant information from the previous hidden state and reset with the current input only. This ability provides a simpler model architecture compared to LSTM. The update gate regulates the flow of information received by the current hidden state from the previous one (Cho et al., 2014). Figure 5.14. shows a simple GRU hidden layer.



Source: Chung et al. (2014)

Figure 5.14: An illustration of GRU

Denote for the j-th GRU hidden unit the candidate activation of the hidden unit as $\hat{h}_j$, the update gate as z_j, and the reset gate as r_j. The actual activation of the proposed unit h_j is then computed by

$$h_j^{(t)} = z_j h_j^{(t-1)} + (1 - z_j)\hat{h}_j^{(t)}, \qquad 5.45$$

where

$$\hat{h}_j^{(t)} = tanh\left([Wx]_j + \left[U\left(r \odot h_{(t-1)}\right)\right]_j\right), \qquad 5.46$$

$$z_j = \sigma\left([W_z x]_j + \left[U_z h_{(t-1)}\right]_j\right), \qquad 5.47$$

$$r_j = \sigma\left([W_r x]_j + \left[U_r h_{(t-1)}\right]_j\right), \qquad 5.48$$

σ is the standard logistic sigmoid function where $\sigma(x) = (1 + e^{-x})^{-1}$ and $tanh$ is the hyperbolic tangent function where $tanh(x) = (e^x - e^{-x})/(e^x + e^{-x})$. x and $h_{(t-1)}$ are the input and the previous hidden state, respectively. W and U are weight matrices which are learned, while $\odot$ is an element-wise multiplication.

Due to its sigmoid activation function, the reset gate forces the hidden state to discard the preceding hidden state when the value is close to zero. On the other hand, the long-term information is learned in the update gate which behaves similarly to the memory cell in LSTM. As each hidden layer has separate reset and update gates, each hidden unit will learn to capture dependencies over different time scales. Therefore, in the case of capturing the short-term dependencies, the reset gates are frequently activated. On the contrary, the update gates are more often active in the event of long-term dependencies (Chung et al., 2014). As in LSTM, the GRUs can be stacked on top of each other as presented in Figure 5.15.

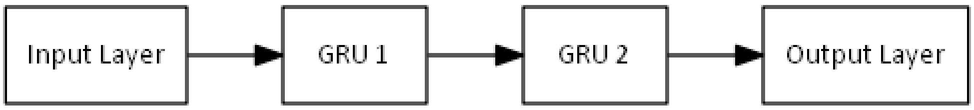


Figure 5.15: An illustration of stacked GRU

5.5.3 Hybrid Model

A typical time series data y_t consists of trend T_t, seasonal S_t, and random variation ε_t that follows the general mathematical form as shown in Equation 5.49 (Prema & Rao, 2015; Montgomery et al., 2015).

$$y_t = f(T_t, S_t, \varepsilon_t) \qquad 5.49$$

There are two common variations that represent the function f depending on the relationship of each component, namely additive and multiplicative models. The additive model sees the time series data as the sum of the component patterns while in the multiplicative model, the data is treated as the product of the component patterns (Prema & Rao, 2015; Montgomery et al., 2015). Equation 5.50 and 5.51 are the formal representation of the additive and multiplicative models, respectively.

$$y_t = T_t + S_t + \varepsilon_t \qquad 5.50$$

$$y_t = T_t \times S_t \times \varepsilon_t \qquad 5.51$$

The rationale behind the hybrid model is that the actual time series often encompasses both linear and nonlinear patterns (Zhang, 2003; Wang et al., 2013). From the perspective of its complexity, the time series data can be composed divided into a linear component and a nonlinear component (Wang et al., 2013). On the other hand, it is difficult to determine if the underlying time series follows a linear or nonlinear process (Zhang, 2003). Moreover, Zhang (2003) proposes a hybrid model combining linear and non-linear elements by using an ARIMA technique to capture the linear part with ANN to handle the non-linear component of the time series data.

Furthermore, following the additive and multiplicative approaches as described previously, analysing the linear and nonlinear component of the time series data can be done by considering additive and multiplicative models as follows:

$$Additive\ model\colon\ y_t = L_t + N_t \tag{5.52}$$

$$Multiplicative\ model\colon\ y_t = L_t \times N_t \tag{5.53}$$

where L_t represents the linear component and N_t the nonlinear component. Put into practice, Zhang (2003) proposes a hybrid model following the additive model while Wang et al. (2013) employs the multiplicative model. Since there is no strong justification for the exact relationship between the linear and nonlinear components of the time series data, it is reasonable for the present study to investigate both additive and multiplicative models on the hybrid model.

In the next section, the design of additive and multiplicative models' implementation on the hybrid model is detailed. This is followed by a dedicated section describing the ANN method that is used to estimate the nonlinear component of the time series data, which is the Nonlinear Autoregressive Neural Network (NARNN).

5.5.3.1. Additive Model

Zhang (2003) considers time series data to be composed of a linear autocorrelation structure and a nonlinear component that follows the additive model as mentioned in Equation 4.52. Both the linear and nonlinear components have to be estimated from the data. First, the ARIMA method is used to model the linear component, so that the residuals from the linear model contain only the nonlinear component. Let r_t denote the residual at time t from the linear model, then

$$r_t = y_t - \hat{L}_t \tag{5.54}$$

where $\hat{L}_t$ is the forecast value for time t from the ARIMA estimation. Zhang (2003) notes that there are currently no general diagnostic statistics to identify nonlinear autocorrelation relationships. Therefore, even if a model has passed diagnostic checking, it may still not be adequate, in that nonlinear relationships may not have been appropriately modelled. Any significant nonlinear pattern in the residuals will indicate the limitation of the ARIMA. Therefore, the residuals from the linear model r_t can be seen as a series of nonlinear components. By modelling residuals using ANNs, nonlinear relationships may be identified. With n input nodes, the ANN model for the residuals is,

$$\hat{r}_t = f\left(r_{t-1}, r_{t-2}, \dots, r_{t-q}\right) + \varepsilon_t \qquad 5.55$$

where f is a nonlinear function determined by the neural network and q is the number of input delays. ε_t is the random error.

Denoting the forecast ANN from Equation 5.55 as $\widehat{N}_t$, the combined forecast is:

$$\hat{y}_t = \widehat{L}_t + \widehat{N}_t \qquad 5.56$$

The additive model's procedures are summarised in Figure 5.16.

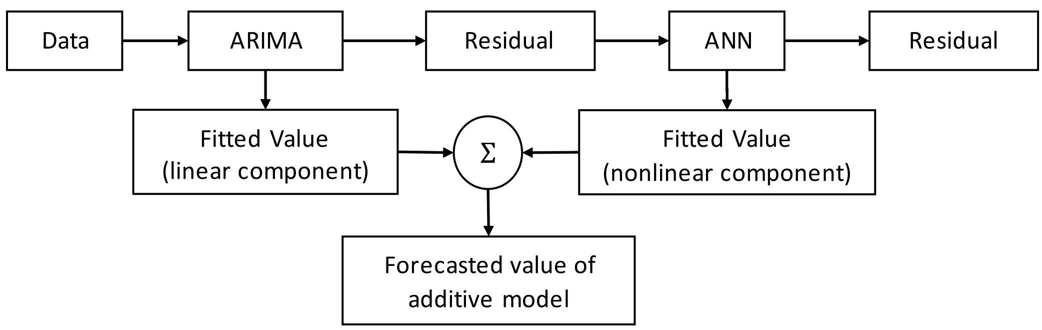


Figure 5.16: The procedures of Additive Model

5.5.3.2. Multiplicative Model

In contrast to Zhang (2003), Wang et al. (2013) consider the relationship between the linear and nonlinear component in a time series data following the multiplicative model as described in Equation 5.53. The first phase of the multiplicative model is the same as the additive model. The linear part of the time series data is estimated using the ARIMA method. As in the additive model, the residuals from the first step contain only the nonlinear component. Therefore, a series of nonlinear components r_t for a multiplicative model is defined by:

$$r_t = \frac{y_t}{\hat{L}_t} \qquad 5.57$$

Where y_t and $\hat{L}_t$ are the actual and forecast value for time t from the ARIMA estimation, respectively.

The next step is estimating the nonlinear components of the time series data by modelling the nonlinear components series r_t using ANN based on Equation 5.55. Once the linear and nonlinear component are estimated, the combined forecast is defined by:

$$\hat{y}_t = \hat{L}_t \times \hat{N}_t \qquad 5.58$$

The multiplicative model's procedures are summarised in Figure 5.17.

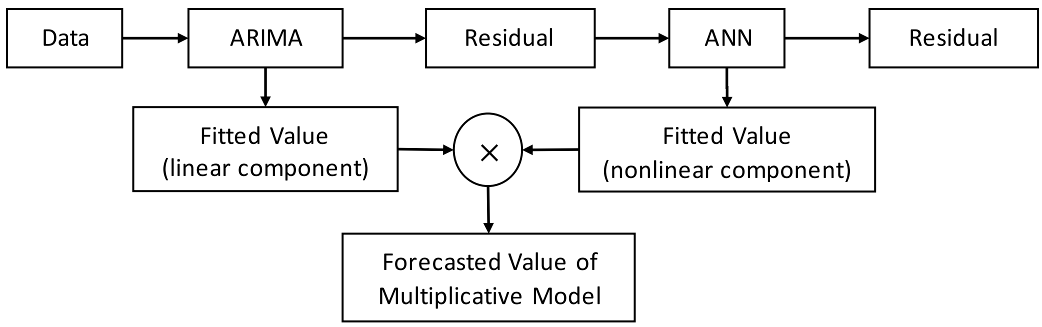
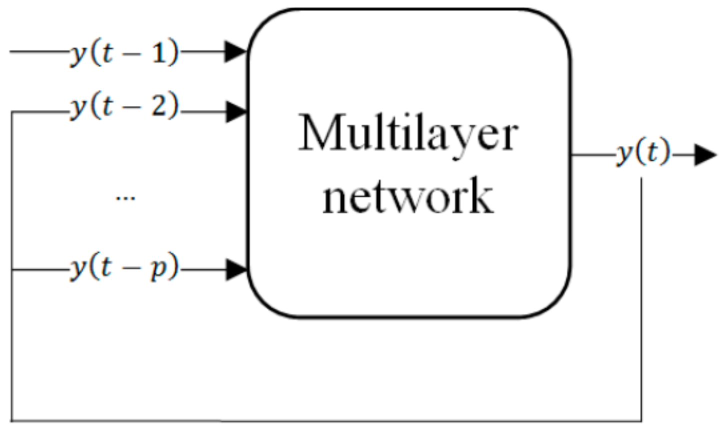


Figure 5.17: The procedures of Multiplicative Model

5.5.3.3. NARNN

The present study uses NARNN to estimate the nonlinear part of the time series data while implementing the hybrid model. The main advantage of implementing NARNN in the case of the hybrid model is its ability to predict a time series data from its past values (Beale et al., 1992). As described in the previous section, the nonlinear part of the time series data is embodied in the residual of the linear model. Therefore, the residual of the ARIMA estimation is the only attribute to capture the nonlinear component of the time series data. Such a condition is ideal for NARNN to be employed as a tool of analysis.

ANN can be classified into static and dynamic neural networks (Zhou et al., 2014). Static neural networks refer to the ANN that produces the output directly from the input via a feed-forward connection without any additional feedbacks or delays, such as MLP and RBF based neural networks. On the other hand, in dynamic neural networks, an extra connection from the previous inputs, outputs, or states of the network is required along with the feed-forward information from input to output. It includes a time-delay network and a recurrent neural network. The NARNN is a dynamic neural network in terms of the use of the previous value of the time series (Zhou et al., 2014). The architecture of NARNN is shown in Figure 5.18.



Source: Ruiz et al. (2016)

Figure 5.18: Architecture of NARNN

The NARNN can be expressed in a mathematical form as follows (Tealab et al., 2018):

$$\hat{y}_t = \sum_{i=1}^{I} w_i f\left(\sum_{j=1}^{p} w_{ij} y_{t-j} + b_i\right) + b_o \qquad 5.59$$

Where $f(.)$ is the activation function, w is the weight parameter of the network, b is the bias parameter, and p is the delays such that $y_{t-1}, y_{t-2}, \cdots, y_{t-p}$ are called feedback delays. Similar to other ANN techniques, the number of hidden layers and the number of neurons per layer are optimised through an iterative, trial-and-error procedure (Ruiz et al., 2016). This study uses the LM algorithm procedure as the learning rule used for the NARNN due to its speed (Ebtehaj & Bonakdari, 2016; Dudek, 2016; Ayala & Coelho, 2016).

5.6. Performance Evaluation Process

The performance of each method is compared to decide which technique is better to model a government cash forecasting. Much research has proposed a range of methods to measure the performance of the forecasting model. However, it is impossible to draw a conclusion of the best forecasting model using only one specific measurement since each performance measurement metric tends to use one type of information while neglecting another perspective of the model (Wallström, 2009). Therefore, this study uses several methods, namely Mean Squared Error (MSE), Root Mean Squared Error (RMSE), Mean Absolute Error (MAE), Mean Absolute Percentage Error (MAPE), Symmetric Mean Absolute Percentage Error (SMAPE), and Mean Absolute Scaled Error (MASE) as the performance evaluation tools.

Moreover, Hyndman and Athanasopoulos (2018) classify the forecasting performance evaluation into three categories, namely (1) scale-dependent errors, (2) percentage errors, and (3) scaled errors. Table 5.2 shows the classification of the proposed performance evaluation based on the three approaches above along with its mathematical representation.

Table 5.2 The classification of the proposed performance evaluation

Performance measurement	Classification	Formula	
MSE	scale-dependent errors	$\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2$	(5.60)
RMSE	scale-dependent errors	$\sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}$	(5.61)
MAE	scale-dependent errors	$\frac{1}{n}\sum_{i=1}^{n}\lvert y_i - \hat{y}_i\rvert$	(5.62)
MAPE	percentage errors	$\frac{1}{n}\sum_{i=1}^{n}\left\lvert\frac{y_i - \hat{y}_i}{y_i}\right\rvert \times 100\ \%$	(5.63)
SMAPE	percentage errors	$\frac{1}{n}\sum_{i=1}^{n}\frac{\lvert y_i - \hat{y}_i\rvert}{(y_i + \hat{y}_i)/2} \times 100\ \%$	(5.64)
MASE	scaled errors	$\frac{1}{n}\sum_{i=1}^{n}\frac{\lvert y_i - \hat{y}_i\rvert}{\lambda}$; $\lambda = \frac{1}{n-1}\sum_{i=2}^{n}\lvert y_i - y_{i-1}\rvert$	(5.65)
where n is the number of observations in the testing set, y_i and $\hat{y}_i$ are actual and predicted value, respectively.			

MSE, RMSE and MAE, the common performance metrics utilised to compare different methods on the same data set, are scale-dependent measures of forecast accuracy (Al-Hnaity & Abbod, 2016; Hyndman & Athanasopoulos, 2018). Although it is relatively easy to compute, MSE, RMSE and MAE have received criticism for their representation values. As the scale-dependent measures, the value of MSE, RMSE and MAE represent the original units of measurement which makes them challenging to qualify whether it is a large error or a small one (Hyndman & Athanasopoulos, 2018; Montgomery et al., 2015). Moreover, Montgomery et al. (2015) argued that the scale-dependent based forecasting performance measurement suffers from a lack of comparability across different time periods. Nevertheless, this study only compares the performance of the forecasting model from different methods at the same time periods. Therefore, the issue is not material in the present study.

As reported by Pontius et al. (2008), some researchers have recommended the use of the Mean Absolute Error (MAE) instead of the RMSE due to its interpretation ability. MAE is easier to understand since it is measuring the average of the absolute values of the errors. Therefore, each error influences MAE in direct proportion to the absolute value of the error, which is not the case for RMSE (Pontius et al., 2008). In addition, the MAE is calculated by taking the absolute value of the difference between the estimated forecast and the actual value, therefore, the negative values do not cancel the positive values (Hyndman & Athanasopoulos, 2018).

An alternative to the scale-dependent based forecasting performance measurement is the percentage error method (Hyndman & Athanasopoulos, 2018). Contrary to the original units-based value of the scale-dependent, the percentage error method of performance metrics measures the relative forecast error of the model (Montgomery et al., 2015). It quantifies the forecast error since the value of the measurement is in percentage form. One can conclude that a 10% error is bigger than a 5% error even when the different data set is employed, for instance. MAPE and SMAPE are considered as the percentage error-based forecasting performance measurement.

MAPE uses the percentage error principle to measure the forecast error with the addition of absolute values of the individual forecast errors (Wallström, 2009). De Myttenaere et al. (2016) argued that the use of MAPE as a loss function for Regression analysis is feasible both from a practical point of view and from a theoretical one since the existence of an optimal model and the consistency of the empirical risk minimisation can be proved.

MAPE is one of the most widely used measures of forecast accuracy, due to its advantages of scale-independency and interpretability. However, MAPE has the significant disadvantage that it produces infinite or undefined values for zero or close-to-zero actual values (Kim & Kim, 2016). SMAPE addresses the shortfall by introducing a modification of MAPE (Hyndman & Athanasopoulos, 2018).

Lastly, MASE is proposed by Hyndman and Koehler (2006) as a scaled error based measurement. It is claimed to be applicable to all modelling scenarios by scaling the error based on the average value of its absolute error. The only circumstance under which these measures would be infinite or undefined is when all historical observations are equal (Hyndman & Koehler, 2006).

For all performance error metrics, the lowest value represents the best model. The proposed performance evaluation techniques are deliberately selected as the representation of each approach proposed by Hyndman and Athanasopoulos (2018). The rationale for such a decision ensures the robustness of the selected government cash forecasting model that satisfies not only from one point of view but also from multiple aspects.

5.7. Summary

This study uses the Indonesian government as the case to investigate the best way to build a government cash forecasting model which focusses on modelling total daily intermittent expenditure of all spending units in Indonesia. Information regarding the government cash disbursement of all spending units in Indonesia is collected from a database owned by the government cash manager of Indonesia. As described in Section 5.3., total daily intermittent expenditure of all spending units in Indonesia is influenced by the total daily available fund for intermittent expenditure, calendar effect, and policy implementation from its respective date. The calendar effect consists of the day of the week, the week of the month, and the month of the year. Therefore, this study utilises total daily intermittent expenditure as predictand and total daily available fund, the day of the week, the week of the month, and the month of the year, and policy implementation as predictors. Total daily intermittent expenditure and total daily available fund for intermittent expenditure are in the form of natural logarithms while the day of the week, the week of the month, and the month of the year, and policy implementation are coded as listed in Table 5.1. The data are from 2009 to 2015 where the data from 2009 to 2013 are used for building the forecasting model while the 2014 to 2015 data are dedicated to testing the performance of the model.

The first stage of this study is the attribute selection process. To ensure the best forecasting model is achievable, the data are split into two datasets based on the variables that are employed to construct the forecasting model. The datasets are called the initial dataset and the attribute selected. In the initial dataset, all variables mentioned previously are utilised to develop a forecasting model while the attribute selected only includes the predictor that significantly influences the value of total daily intermittent expenditure. This study takes advantage of the correlation value of each variable relative to the predictand.

The next stage is the modelling process where the development of the forecasting model is happening. This study uses three types of method to construct the forecasting model which are ARIMA, ANN, and hybrid model. The method selection is intended to investigate the characteristics of the data. It is argued that there are three different patterns which real-world time series data may follow: linear, nonlinear, or both linear and nonlinear patterns. Each selected method is inherited with the ability to capture a specific time series pattern. ARIMA model is capable of explaining the linearity embodied in a time series data while ANN is better in describing the nonlinearity of the data. In the case where both linear and nonlinear patterns are existing in a time series data, the hybrid model is the best method to confirm it.

Furthermore, ARIMAX is selected as the modelling technique of the ARIMA method due to a practical reason. This research bases its study on the assumption that the value of total daily intermittent expenditure is influenced by exogenous variables described in Section 5.3. Multiple ANN techniques are proposed in this study to explore the best method to develop a government forecasting model. The proposed ANN techniques, namely FFNN, CFNN, RBFN, GRNN, ELM, LSTM, and GRU, are selected as a representation of most ANN classifications. This study combines ARIMAX and NARNN into a hybrid model. Both additive and multiplicative models are investigated to capture the linear and nonlinear relationship of the observed time series data.

Lastly, the best government cash forecasting model with accuracy that meets an acceptable level of materiality for the cash manager is selected by comparing the performance of each model in the performance evaluation phase. Most forecasting performance evaluation can be classified into three categories based on its measurement approach. Each approach measures a particular performance aspect of the model. This study used six metrics namely MSE, RMSE, MAE, MAPE, SMAPE, and MASE to ensure the best selected forecasting model was agreed by most performance criteria. The experiments conducted in this study followed the procedure above. The next chapter reports the results and discusses the findings.

Chapter 6 Experiments and Discussions

6.1. Introduction

The purpose of this chapter is to discuss the results of the experiments with respect to identifying the best forecasting model with the accuracy that meets an acceptable level of materiality for the cash manager. Overall, the experiments conducted in this research can be broken down into two parts based on the variables that are used in the models. The first part employed all variables explained in Section 5.4. namely the Initial Dataset, while the other part utilised only the relevant variables chosen through an attribute selection process called the Selected Attributes. For both the Initial Dataset and Selected Attributes, the proposed method as described in Section 5.5. was used to develop a forecasting model. The performance of the models was compared to identify the best model to accurately predict government cash forecasting based on the measurements introduced in Section 5.6.

The chapter is organised as follows. The attribute selection process is presented in Section 6.2. The models proposed by ARIMA, ANN, and Hybrid techniques are detailed in Section 6.3. The results and discussion on the dataset that was used for each method are detailed with the model built on the initial dataset presented first, followed by the model developed on the selected attributes. Analysis of the proposed models and a rationale for the models selected are presented in Section 6.4. Lastly, results and discussions are summarised in Section 6.5.

6.2. Attribute Selection

There are different opinions on the implementation of the attribute selection process before the modelling phase. Some argue, e.g. Guyon & Elisseeff (2003) that it helps to gain a simpler model and improved forecasting ability while others, e.g. Heinze et al. (2018) suggest the opposite by saying a variable selection process leads to some problems such as robustness and interpretation of the models. In finding the best forecasting model, this study took advantage of the initial dataset which consisted of all the related variables available and the selected attributes from the variable selection process. The detailed information regarding both datasets is described in the next two sections.

6.2.1 Initial Dataset

The initial dataset took all the variables described in Section 5.3. to develop a government cash forecasting model. It consisted of one dependent variable, which was total daily intermittent expenditure, E, and five independent variables which were total funds available daily for intermittent expenditure, F, the day of the week, D, the week of the month, W, the month of the year, M, and policy implementation, $Policy$. As listed in Table 5.1, variable E and F were continuous data while variable D, W, M, and $Policy$ were categorical. For modelling purposes, all categorical variables were transformed into binary variables and represented each category in the variables. For an example, variable D was transformed into five variables such that variable $D1$ indicates that the data occurred on Monday only, while variable $D2$ represents the information which happened only on Thursday, and so forth. The binary variable can only take two values, “1” if the category existed in the data and “0” if it did not. Therefore, the value of variable $D1$ is “1” for each observation that happened on Monday and “0” for any other day. In total, there were 25 attributes treated as independent variables in the Initial Dataset. The complete list of the initial dataset is presented in Appendix 1.

6.2.2 Selected Attributes

As described in Chapter 5, this study used the CAE method to select the predictor variables that significantly influence the independent variable. The method was implemented in the WEKA 3.8 software environment. It allowed this study to calculate the correlation between categorical/nominal and continuous/numeric data by automatically converting the categorical attributes into its binary indicators before giving the result of overall correlation as a weighted average value (Hall et al., 2009; Frank et al., 2016). The CAE method ranked the value of Pearson’s correlation value between all independent variables used in this study to the total daily intermittent expenditure, E, as shown in Table 6.1.

Table 6.1: Correlation Attribute Evaluation values

Value	Variable	Sign
0.2537	F	Negative
0.1138	M	Positive
0.0730	$Policy$	Positive
0.0639	W	Positive
0.0222	D	Positive

Based on the value of CAE, the total fund available daily for intermittent expenditure, F, is the most significant variable. The negative sign confirms the relationship between the total daily available fund for intermittent expenditure, F, and the total daily intermittent expenditure, E, where a decrease in the total daily available fund for intermittent expenditure will increase the total daily intermittent expenditure. The day of the week, D, the week of the month, W, and the month of the year, M, which are the proxy to calendar effect, are in a positive correlation. These results imply the total daily intermittent expenditure was increasing along the budget year. The results are understandable when referring to the graph of cash disbursement patterns in Figure 3.1.

Figure 3.1 shows the percentages of quarterly total intermittent expenditure for all spending units in Indonesia. The disbursement activity of spending units tended to be low in the first quarter of the budget year followed by a moderate increase in the second and third quarter. The increase continued in the last quarter of the budget year. On average, the increase in the total daily intermittent expenditure was doubled during the last quarter compared to the previous one.

As described in Section 3.4., the policy implementation, $Policy$, is the proxy for the implementation of new policy on government cash disbursement. The CAE result showed a positive correlation between $Policy$ and E, which means the new policy introduced by the cash manager intensified the total daily intermittent expenditure. In the case of this study, the new policy deployed in 2011 increased the total daily intermittent expenditure which was in line with the purpose of the policy itself.

However, a further discussion on variables correlation is beyond the focus of this study.

The lowest value of CAE was the day of the week, D, that recommended the variable did not significantly influence the total daily intermittent expenditure with the value of 0.0222. It was also supported by the fact that the value of CAE of the day of the week, D, dropped expressively at 65.26%, comparing the CAE's value to the week of the month, W, at 0.0639 on the second last rank. This figure, strongly suggested that this study should exclude the day of the week, D, variable from the dataset and assembling the new dataset called the selected attributes which consisted of one dependent variable which was total daily intermittent expenditure, E, and four independent variables which were total daily available fund for intermittent expenditure, F, the week of the month, W, the month of the year, M, and policy implementation, $Policy$.

To support this selection, several datasets gathered based on the rank of the CAE's value were employed to develop a series of forecasting models. The first dataset was the one with all the proposed variables. The next dataset was generated from the first dataset with an exception of the less correlated variables. The routine was repeated up to the last dataset which contained only the highest score of CAE. The comparison of the performance of all datasets agreed with the previous selection where total daily available fund for intermittent expenditure, F, the week of the month, W, the month of the year, M, and policy implementation, $Policy$ were selected. The detailed results can be found in Appendix 2.

Like the Initial Dataset, all categorical variables in the selected attributes were transformed into binary variables representing each category in the variables. The transformation led the selected attributes dataset to have one dependent variable and 20 independent variables as listed in Appendix 3.

6.3. Modelling

In this section, the government cash forecasting model that was developed using the proposed methods and datasets is presented. The parameters of each model were designated such that the best performance was achievable. Multiple software environments such as EViews, MATLAB, and Python, were used as an implementation tool of the methods.

6.3.1 Autoregressive Integrated Moving Average Model

This study proposed several variables that affect the total daily intermittent expenditure, E. To accommodate them, a branch of the ARIMA model so called ARIMA with Exogenous Variables (ARIMAX) was used to develop a government cash forecasting model. The method was developed under the EViews 9.5.

As a statistical method, ARIMA technique requires the estimated time series data to be stationary. Therefore, the unit root test was assigned to total daily intermittent expenditure, E and total daily available fund for intermittent expenditure, F. Furthermore, as described in Section 5.5.1., when implementing the ARIMA method, ensuring the stationarity of the variables is not only a matter of having a robust estimation and the generalisation of the produced forecasting model, but also essential for the model's configuration. The level of integration (d) of the dependent variable, which was necessary to construct an ARIMA (p, d, q) model, was determined from the result of the unit root test.

This study utilised an Augmented Dickey-Fuller (ADF) unit root test to define the level of integration of the time series variables. As shown in Table 6.2., the unit root test results indicated that all of the time series variables used in this study were stationary or integrated at level 0/$I(0)$. Hence, the model constructed using these variables was free from the potential problem mentioned above.

Table 6.2: Unit root test of variable E and F

Variable	ADF t-statistics	Test critical values			Conclusion
		1%	5%	10%	
$\boldsymbol{E}$	-13.38	-3.44	-2.86	-2.57	Integrated at level 0 /$I(0)$
$\boldsymbol{F}$	-3.85				Integrated at level 0 /$I(0)$

Moreover, the level of integration (d) for ARIMAX(p,d,q) could be definite from the level of integration of total daily intermittent expenditure, E. Therefore, the level of integration (d) for the ARIMAX(p,d,q) was 0.

Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
		1	0.679	0.679	567.83	0.000
		2	0.556	0.175	948.31	0.000
		3	0.461	0.057	1209.7	0.000
		4	0.317	-0.105	1333.8	0.000
		5	0.242	-0.001	1405.9	0.000
		6	0.206	0.052	1458.2	0.000
		7	0.161	0.011	1490.4	0.000
		8	0.131	-0.003	1511.6	0.000
		9	0.105	-0.012	1525.2	0.000
		10	0.103	0.041	1538.4	0.000
		11	0.101	0.028	1551.0	0.000
		12	0.081	-0.019	1559.2	0.000
		13	0.079	0.007	1566.9	0.000
		14	0.059	-0.021	1571.2	0.000
		15	0.067	0.045	1576.8	0.000
		16	0.065	0.007	1582.0	0.000
		17	0.055	-0.009	1585.7	0.000
		18	0.051	-0.004	1589.0	0.000
		19	0.047	0.007	1591.8	0.000
		20	0.047	0.018	1594.5	0.000
		21	0.043	-0.003	1596.9	0.000
		22	0.041	0.001	1599.0	0.000
		23	0.041	0.005	1601.2	0.000
		24	0.038	0.006	1603.0	0.000
		25	0.033	-0.003	1604.3	0.000
		26	0.037	0.011	1606.0	0.000
		27	0.039	0.010	1607.9	0.000
		28	0.029	-0.014	1609.0	0.000
		29	0.031	0.008	1610.2	0.000
		30	0.028	0.000	1611.2	0.000
		31	0.028	0.008	1612.1	0.000
		32	0.029	0.005	1613.2	0.000
		33	0.029	0.001	1614.2	0.000
		34	0.023	-0.008	1614.9	0.000
		35	0.016	-0.008	1615.2	0.000
		36	0.012	0.001	1615.4	0.000

Figure 6.1: Correlogram of the ACF and PACF for E

Once the level of integration, (d) was known, the next step in developing the government cash forecasting model using ARIMA was to determine the value of maximum lag lengths for the AR and MA terms. This was done by analysing the correlograms of the Autocorrelation Function (ACF) and Partial ACF of the series. As shown in Figure 6.1., maximum lag lengths suggested for the AR and MA terms was

4. Therefore, the proposed ARIMAX(p, d, q) architecture was the one that had an autoregressive term (p) of maximum 4, level of integration (d) of 0, and moving average term (q) of maximum 4.

The next step was the modelling stage which included investigating the best ARIMAX(p, d, q) model. It was done by estimating the data following all possible configurations. The best ARIMA specification was chosen based on several model selection criteria which were the Akaike Information Criterion (AIC), Bayesian Information Criterion (BIC), and Hannan-Quinn Information Criterion (HQC). The smaller value of model selection criteria represented the better model to choose. The results and discussion on the ARIMAX modelling process using the initial dataset and the selected attributes are presented in the following section.

Initial dataset

Table 6.3: Model Selection Criteria for initial dataset

Model	AIC	BIC	HQC	Model	AIC	BIC	HQC
(2,0,3)	**3.677759**	**3.794421**	**3.721656**	(1,0,2)	3.696006	3.804335	3.736768
(4,0,1)	3.690588	3.807251	3.734486	(2,0,4)	3.696537	3.817366	3.742002
(3,0,1)	3.691184	3.803680	3.733514	(4,0,0)	3.699856	3.812352	3.742185
(3,0,2)	3.691209	3.807871	3.735106	(3,0,0)	3.716584	3.824913	3.757346
(4,0,2)	3.691940	3.812769	3.737405	(2,0,0)	3.717758	3.821921	3.756952
(4,0,4)	3.692464	3.821626	3.741065	(2,0,1)	3.717915	3.826245	3.758677
(3,0,4)	3.692525	3.817521	3.739558	(1,0,1)	3.748344	3.852507	3.787538
(2,0,2)	3.693410	3.805907	3.735740	(1,0,0)	3.842477	3.942474	3.880104
(3,0,3)	3.694140	3.814969	3.739605	(0,0,4)	3.945831	4.058328	3.988161
(4,0,3)	3.694421	3.819417	3.741454	(0,0,3)	4.013011	4.121341	4.053773
(1,0,3)	3.694571	3.807067	3.736901	(0,0,2)	4.135546	4.239709	4.174741
(1,0,4)	3.694816	3.811479	3.738714	(0,0,1)	4.249321	4.349317	4.286947
				(0,0,0)	4.517391	4.613221	4.553450

Once all possible ARIMAX specifications were estimated, the value of model selection criteria for each ARIMAX model was then compared. The best ARIMAX model was selected based on the smallest model selection criteria value. Table 6.3. listed the value of model selection criteria for all possible ARIMAX specifications. All model selection criteria agreed that ARIMA(2,0,3) had the best specifications to model a government cash forecasting using initial data which provided the performance scores of 1.380, 1.1749, 0.7252, 2.7862, 2,915, and 1.336 when it was measured with MSE, RMSE, MAE, MAPE, SMAPE, and MASE respectively.

<u>Selected Attributes</u>

The same model selection procedures were applied for the forecasting model that employed the selected attributes only. However, as shown in Table 6.4., the best ARIMAX model was inconclusive. AIC value suggested that the best model was ARIMAX (4,0,4), while BIC and HQC agreed on ARIMA(1,0,2). Therefore, the next examination was necessary to define the best ARIMA model to predict government cash needed in the future.

Table 6.4: Model Selection Criteria for selected attributes

Model	AIC	BIC	HQC	Model	AIC	BIC	HQC
(4,0,4)	**3.723509**	3.802673	3.753297	(4,0,1)	3.728838	3.795503	3.753923
(1,0,2)	3.725059	**3.783390**	**3.747008**	(4,0,0)	3.729078	3.791576	3.752594
(1,0,3)	3.725395	3.787893	3.748912	(3,0,1)	3.733507	3.796004	3.757023
(2,0,2)	3.725410	3.787908	3.748927	(3,0,0)	3.745305	3.803637	3.767254
(2,0,3)	3.727025	3.793690	3.752110	(2,0,1)	3.748790	3.807121	3.770739
(1,0,4)	3.727025	3.793690	3.752110	(2,0,0)	3.751030	3.805195	3.771411
(3,0,2)	3.727027	3.793692	3.752112	(1,0,1)	3.783425	3.837590	3.803806
(2,0,4)	3.727509	3.798340	3.754161	(1,0,0)	3.868244	3.918242	3.887057
(3,0,3)	3.728074	3.798905	3.754726	(0,0,4)	4.206400	4.268898	4.229916
(4,0,3)	3.728228	3.803226	3.756448	(0,0,3)	4.319672	4.378004	4.341621
(3,0,4)	3.728530	3.803527	3.756750	(0,0,2)	4.474224	4.528389	4.494605
(4,0,2)	3.728617	3.799448	3.755269	(0,0,1)	4.641832	4.691830	4.660645
				(0,0,0)	5.042379	5.088211	5.059625

Since the main objective of this study was to develop a government cash forecasting model with the accuracy that meets an acceptable level of materiality for the cash manager, the best model proposed had to be the one with the best performance accuracy. Therefore, the best model was decided by examining the performance accuracy of each model recommended by model selection criteria. The comparison of the performance between ARIMAX (4,0,4) and ARIMAX (1,0,2) is presented in Table 6.5. It is clear from the table that ARIMAX (1,0,2) is the best model to forecast the forthcoming cash requirement of a government. Furthermore, ARIMA (1,0,2) provided the performance scores of 1.2903, 1.1359, 0.6059, 2.3662, 2.5480, and 1.1166 when it was measured with MSE, RMSE, MAE, MAPE, SMAPE, and MASE respectively.

Table 6.5: Performance comparison of ARIMAX (4,0,4) and ARIMAX (1,0,2)

Model	MSE	RMSE	MAE	MAPE	SMAPE	MASE
ARIMAX (4,0,4)	1.4083	1.1867	0.6449	2.51091	2.7225	1.1883
ARIMAX (1,0,2)	**1.2903**	**1.1359**	**0.6059**	**2.3662**	**2.5480**	**1.1166**

The experiment above shows that utilising only the significant variables that are chosen via the attribute selection process could raise the performance of the ARIMAX models. Figure 6.2. displays the residual of the proposed ARIMAX model for both using initial dataset and selected attributes.

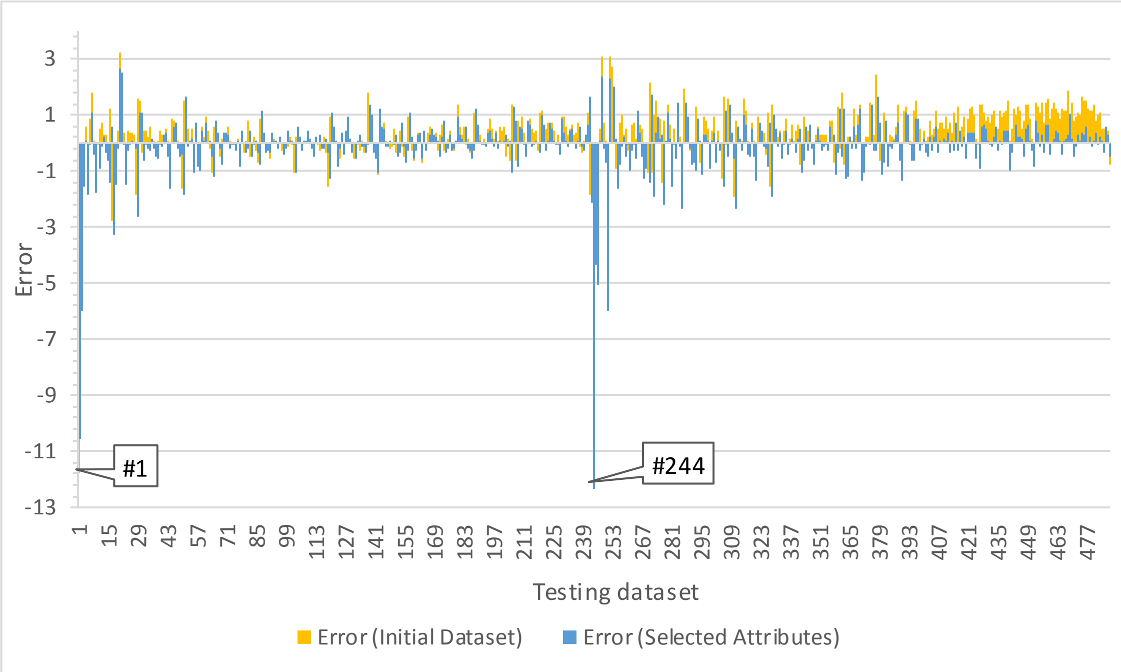


Figure 6.2: The residual of the ARIMAX model

As mentioned in Section 5.3., the testing dataset spanned the period of 2014 – 2015 to assess the ability of the models capturing the shift between budget years. Therefore, Observations 1 and 244 in Figure 6.2. represent the first working day in the 2014 and 2015 budgets respectively. The values of residuals for both models are the positive and negative signs. The signs represent the quality of the forecasted value of the model where the positive sign indicates over-forecasting and a negative sign indicates under-forecasting. Over-forecasting happens when the forecasted value lies above the actual value while the one that falls below the actual value is considered under-forecasting. However, whichever direction the deviation of the forecasted values is relative to its actual, the absolute values of the error determine the performance of the model. The residuals with the value of 0.5 and -0.5, for instance, are equally worthy. The same arrangement was applied for the upcoming figure 6.5, 6.8, 6.11, 6.14, 6.17, 6.20, 6.23, 6.28 that presents the residual of the proposed government cash forecasting models.

It is obvious from Figure 6.2. that some large deviations occurred throughout the testing dataset. The outliers occurred during the first quarter of the budget year (e.g. Observations 1, 2, 17, 20, 244, 246, 248, 251) where the spending units' expenditures were relatively low compared to the other quarters of the budget year as presented in Figure 6.2. As an illustration, the biggest deviation observed in the testing dataset number 244 was on the 2nd of January 2015. The absolute value of the error from both models on the particular day is 9.7962 and 12.3134 for the model that was developed using the initial dataset and selected attributes respectively. It implied that the most noticeable feature of the figure suggests the use of the initial dataset as a better technique to model the future of the spending units' expenditure compared to the selected attributes. However, another obvious figure suggested the opposite. Observation 1 which was on the 2nd of January 2014 provided the value of 11.7424 and 10.5277 for the absolute value of the error from the models developed using the initial dataset and selected attributes respectively which supports the selected attributes as a better dataset to develop a government cash forecasting model. Both observations illustrated the variation of results throughout the testing dataset. The errors from both models fluctuate and tend to overlap each other along the observation period. Further investigation recommended most of the residuals of the ARIMAX model developed from the initial dataset were larger than the errors of the model from the selected attributes. It confirms the previous conclusion derived from the performance measurements of the models described earlier, that the ARIMAX model performs better when utilising the selected attributes data compared to the one with the initial dataset.

Furthermore, Figure 6.2. also demonstrates that both models had difficulty to forecast the expenditure of spending units during the beginning of the budget year where the large deviation between the actual and the forecasted value of the expenditure only happens during the first month of each budget year. The proposition for such a condition can be traced back to the disbursement patterns of the spending units. As explained in Section 3.4., the nature of the type of expenditure influences the spending behaviour of the government agencies in Indonesia. The spending units' construction projects may be concluded at the end of the budget year. Some expenditures, such as capital expenditure, are embodied in the specific procurement process that may delay the cash disbursement of the spending units. The accumulated spending toward the end of the budget year causes low disbursement at the beginning of the budget year. Improvement in the procurement process and disbursement mechanism might stimulate the spending behaviour. However, the discussion in this regard is outside of the focus of the present study. Despite its limitation, the proposed ARIMAX models are capable of producing a reasonable accuracy when measured with the proposed performance measurement evaluation tools.

6.3.2 Artificial Neural Network Model

This study used multiple ANN techniques to investigate the best procedure for developing a government cash forecasting model with acceptable accuracy for the cash manager. FFNN, CFNN, RBFNN, GRNN, and ELM were implemented using MATLAB R2018b software while Python 3.6.5 was used to execute LSTM and GRU techniques. The training dataset was used to develop a forecasting model. A trial-and-error approach was applied to construct each ANN architecture with the goal of finding the most effective structure that produces the most accurate government cash forecasting model. It is an iterative process where the best architecture is chosen such that increasing or decreasing the ANN parameters ultimately fails to increase the model's performance. The testing dataset was used to measure the model's performance following several evaluation techniques. The following sections discuss the results of each ANN technique proposed in this study as described in Section 5.5.2.

6.3.2.1. Feed-forward Neural Network Model

The FFNN structure consists of one input layer, hidden layers, and one output layer. This study used hyperbolic tangent sigmoid function (tansig) as activation function of hidden layers and linear transfer function (purelin) for the output layer which is default to feedforward network function in MATLAB (Beale et al., 1992). Moreover, the Levenberg-Marquardt Algorithm (trainlm) was chosen as the training function. Therefore, the parameters that need to be set to construct a forecasting model are the number of hidden layers and the number of neurons in each hidden layer.

Initial dataset

This study proposed an FFNN network with three hidden layers with three neurons in each layer as the best model when the initial dataset was used. The chosen FFNN architecture can be seen in Figure 6.3. The network gave performance scores of 0.581, 0.762, 0.509, 1.900, 1,911, and 0.938 when it was measured with MSE, RMSE, MAE, MAPE, SMAPE, and MASE respectively.

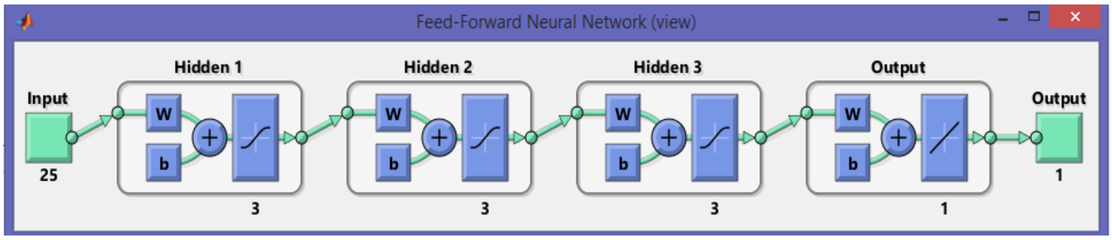


Figure 6.3: FFNN Architecture for initial dataset

Selected Attributes

In the case where the selected attributes were employed, the FFNN method proposed a network with three hidden layers with five neurons in each layer as the best model. The network gave performance scores of 0.544, 0.737, 0.506, 1.893, 1,909, and 0.933 when it was measured with MSE, RMSE, MAE, MAPE, SMAPE, and MASE respectively. FFNN Architecture for selected attributes is shown in Figure 6.4.

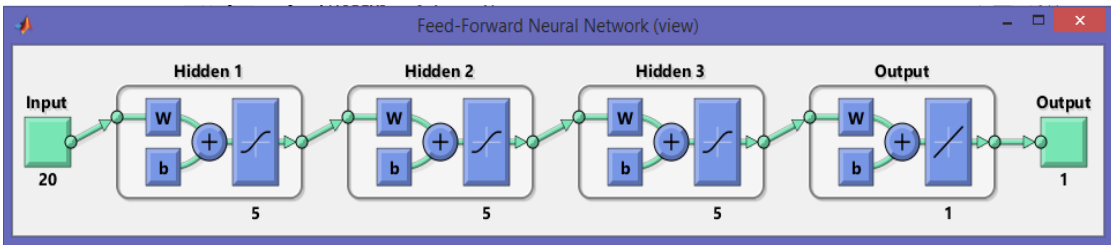


Figure 6.4: FFNN Architecture for selected attributes

The results showed that the performance of the FFNN model to predict government cash demand in the future could increase by including only the significant attributes. The residual of the proposed FFNN model for both using initial dataset and selected attributes are presented in Figure 6.5.

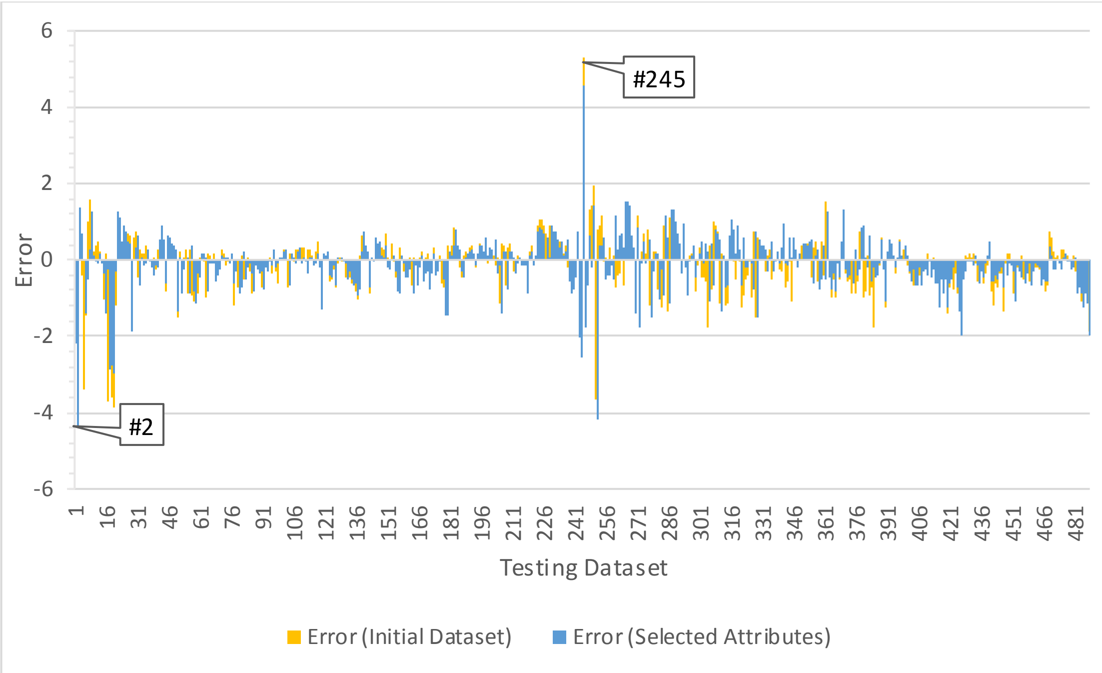


Figure 6.5: The residual of the FFNN model

The X-axis and Y-axis of Figure 6.5. are identical to the ones described in Figure 6.2. Figure 6.5. shows some large deviations occurred throughout the testing dataset. The outliers occurred during the first quarter of the budget year (e.g. Observations 2, 17, 18, 19, 245, 251) where the spending units' expenditures were relatively low compared to the other quarters of the budget year as presented in Figure 6.5. As an illustration, the biggest error observed in the testing dataset number 245 was on the 5th of January 2015. The absolute value of the error from both models on the particular day was 5.3058 and 4.5598 for the model that was developed using the initial dataset and selected attributes respectively. The most obvious feature of the figure suggests the superiority of the model that was developed using the selected attributes over the one that was built under the initial dataset to project the government cash needed in the future. However, another obvious figure suggested the opposite. Observation 2 which was on the 3rd of January 2014 provided the value of 3.4677 and 4.40682 for the absolute value of the error from the models developed using the initial dataset and selected attributes respectively. The values indicate the initial dataset as a better dataset to develop a government cash forecasting model compared to the selected attributes. Both observations presented the diversity of results throughout the testing dataset. Although both models produce deviations that are oscillated and overlapped each other along the observation period, Figure 6.5. supports the idea of the selected attributes as the best dataset to develop a government cash forecasting model. Most of the errors from the initial dataset were larger than the errors of the model that used the selected attributes, when the FFNN method was utilised to develop the model. The inference agrees with the results from performance measurements as discussed earlier that the best model proposed by the FFNN method is achieved when utilising the selected attributes data.

Similarly to the previous method, Figure 6.5. also reveals that predicting the expenditure of spending units during the beginning of the budget year is a challenging task. The forecasted expenditure of spending units from the models developed under the FFNN technique deviated in a large margin compared to its actual value. It only happened during the first month of each budget year. Investigating the disbursement patterns of the spending units might explain the phenomenon. The preceding Section 3.4. explained that the nature of the type of expenditure plays a significant role to the spending behaviour of the government agencies in Indonesia. In most cases, the spending units' construction projects are concluded at the end of the budget year. Some expenditures, such as capital expenditure, are embodied in the specific procurement process that delays the cash disbursement of the spending units up to the later part of the budget year. The accumulated spending toward the end of the budget year causes the low disbursement at the beginning of the budget year which makes the predicting ability of the model on the respective period problematic. Improvement on the procurement and disbursement policy might stimulate the spending behaviour. However, the discussion in this regard is outside of the focus of the present study. Despite its limitation, the proposed FFNN models are capable of providing a reasonable accuracy when they are measured with the proposed performance measurement evaluation tools.

6.3.2.2. Cascade-forward Neural Network Model

The structure of CFNN has a similarity to feedforward networks except for extra connections from all of its previous layers such that the second hidden layer will have not only directed connections from the first hidden layer but also direct connections from the input layer, for instance (Beale et al., 1992). This study used the default activation function for both hidden layers and the output layer and the Levenberg-Marquardt Algorithm as the training function. Similarly to FFNN, the number of hidden layers and the number of neurons in each hidden layer are to be set by the user in order to construct a forecasting model.

<u>Initial dataset</u>

This study proposed a CFNN network with three hidden layers with three neurons in each layer as the best model. The chosen CFNN architecture can be seen in Figure 6.6. The network gave performance scores of 0.618, 0.786, 0.518, 1.979, 1,960, and 0.954 when it was measured with MSE, RMSE, MAE, MAPE, SMAPE, and MASE respectively.

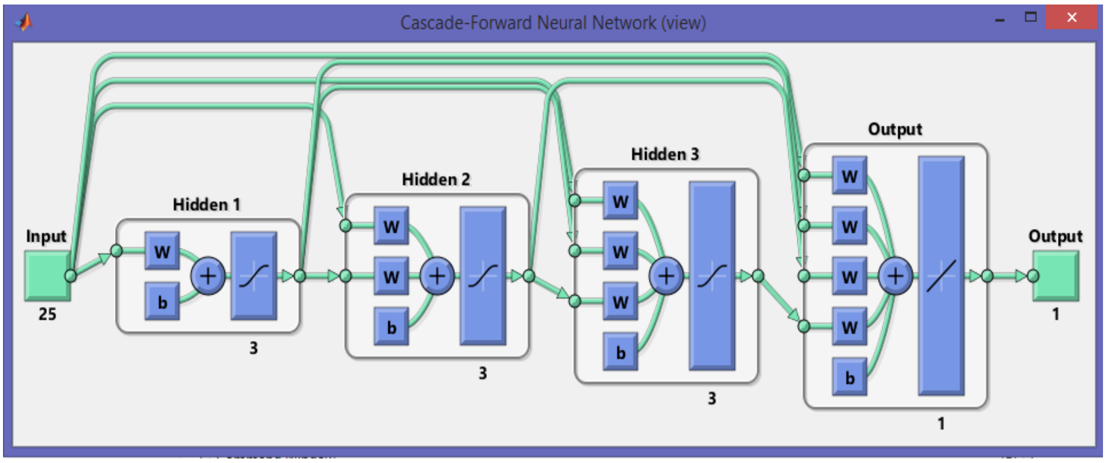


Figure 6.6: CFNN Architecture for initial dataset

<u>Selected Attributes</u>

The CFNN method proposed a network with two hidden layers with three neurons in each layer as the best model with the selected attributes. The network gave performance scores of 0.567, 0.753, 0.505, 1.915, 1,910, and 0.931 when it was measured with MSE, RMSE, MAE, MAPE, SMAPE, and MASE respectively. FFNN Architecture for selected attributes is shown in Figure 6.7.

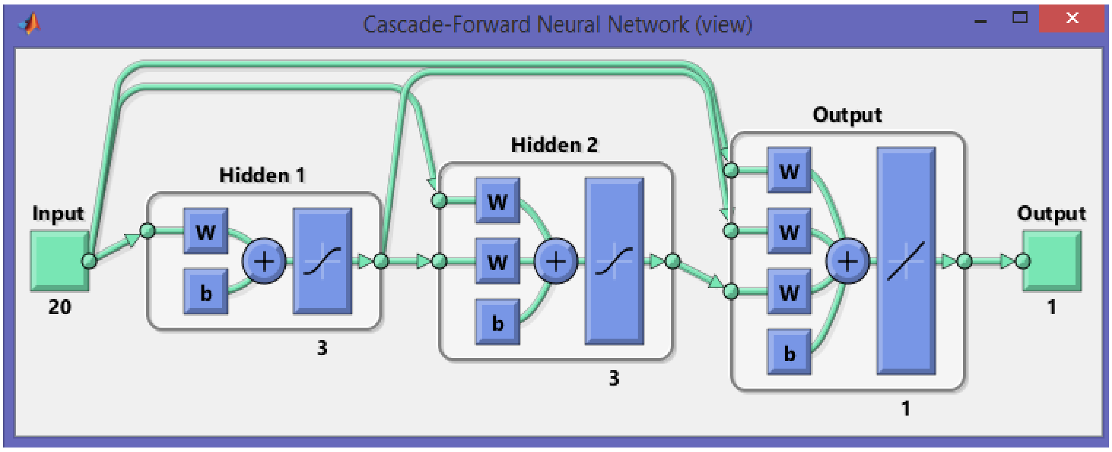


Figure 6.7: CFNN Architecture for selected attributes

When both results are compared, it provides evidence that the attribute selection process increases the performance of the CFNN model to predict the forthcoming government cash requirement. The residual of the proposed CFNN model for both using initial dataset and selected attributes is presented in Figure 6.8.

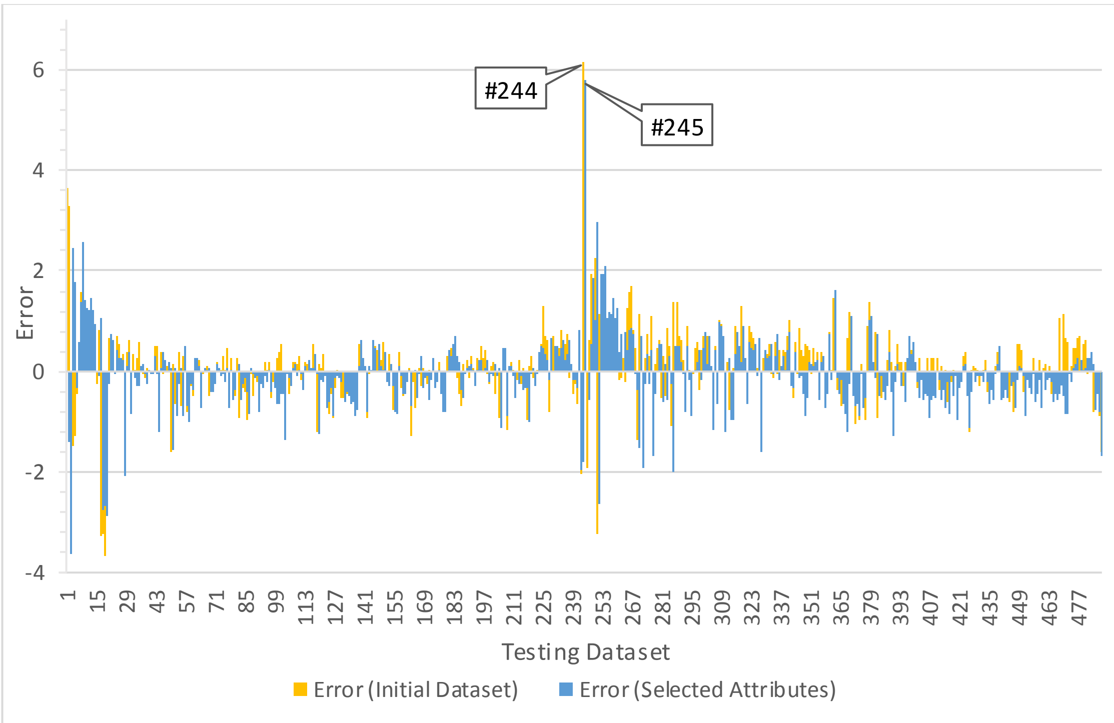


Figure 6.8: Figure The residual of the CFNN model

The vertical and horizontal axes of Figure 6.8. are identical to the ones described in Figure 6.2. Some large deviations between the forecasted and actual value during particular days throughout the testing dataset are observed in Figure 6.8. The anomaly occurred during the first quarter of the budget year (e.g. Observations 1, 2, 17, 19, 244, 245, 250) where the spending units' expenditures were relatively low compared to the other quarters of the budget year as presented in Figure 6.8. As an illustration, the biggest deviation observed in the testing dataset number 244 is on the 2nd of January 2015. The absolute value of the error from both models on the particular day is 6.1385 and 1.7937 for the model that was developed using the initial dataset and selected attributes respectively. The most noticeable feature of the figure suggests the superiority of the model that was developed using the selected attributes over the one that was built under the initial dataset to project the government cash needed in the future. Nevertheless, another obvious figure suggested the contrary. Observation 245 which was on the 5th of January 2015 providing the values of 3.211748 and 5.780084 for the absolute value of the error from the models developed using the initial dataset and selected attributes respectively. It indicates that in comparison with the selected attributes, the initial dataset is a better dataset to develop a government cash forecasting model. The fluctuation of results throughout the testing dataset is illustrated by both observations. Despite of the fact that the errors from both models are varied and tended to overlap each other along the observation period, Figure 6.8. supports the suggestion that the best forecasting model developed utilising the CFNN technique can be accomplished by using the selected attributes dataset. This was concluded from the fact that most of the errors of the model developed from the initial dataset were larger than the selected attributes, when the CFNN method was used. It also confirms the previous conclusion derived from the performance measurements that by utilising the selected attributes data, the CFNN model is superior compared to the one with the initial dataset.

Figure 6.8. also suggests that regardless the employed data, it was problematic for both models to forecast the expenditure of spending units during the beginning of the budget year. The large deviation between the actual and the forecasted value of the expenditure only occurred during the first month of each budget year. The possible cause for such a condition can be traced back to the disbursement patterns of the spending units. As explained in Section 3.4., the nature of the type of

expenditure influences the spending behaviour of the government agencies in Indonesia. The spending units' construction projects may be concluded at the end of the budget year. Some expenditure, such as capital expenditure, is embodied in the specific procurement process that may delay the cash disbursement of the spending units. The aggregation of spending units' expenditure at the end of the budget year causes low disbursement at the beginning of the budget year. Improvement in the procurement process and disbursement mechanism might stimulate the spending behaviour. However, the discussion in this regard is outside of the focus of the present study. Despite its limitation, the proposed CFNN models are capable of producing a reasonable accuracy when they are measured with the proposed performance measurement evaluation tools.

6.3.2.3. Radial Basis Function Neural Network Model

RBFNN consists of two layers which are a hidden layer and an output layer. The most important distinction between RBFNN and other ANN architectures is noticeable from its activation function. RBFNN uses a radial basis function in the hidden layer while using the linear activation function for its output layer. In MATLAB, the spread of radial basis functions and the number of neurons in the hidden layer are determined by the user.

Initial dataset

This study proposed an RBFNN network with nine spreads with a hundred neurons as the best model with the initial dataset. The network gave performance scores of 1.130, 1.063, 0.751, 2.781, 2.811, and 1.385 when it was measured with MSE, RMSE, MAE, MAPE, SMAPE, and MASE respectively. RBFNN architecture for the initial dataset is shown in Figure 6.9.

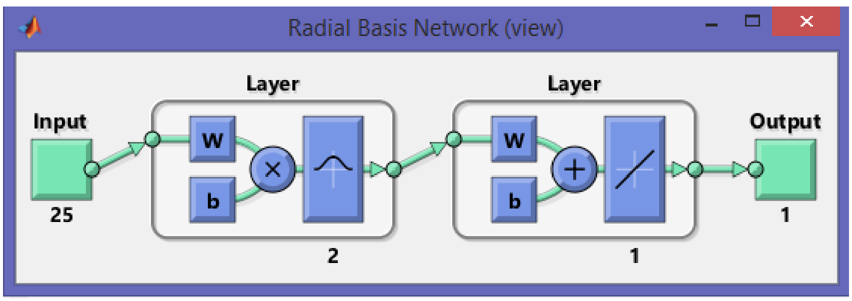


Figure 6.9: RBFNN Architecture for initial dataset

Selected Attributes

This study proposed an RBFNN network with one spread with a hundred neurons as the best model with the selected attributes. The network gave performance scores of 1.025, 1.012, 0.697, 2.591, 2.612, and 1.285 when it was measured with MSE, RMSE, MAE, MAPE, SMAPE, and MASE respectively. RBFNN Architecture for selected attributes is shown in Figure 6.10.

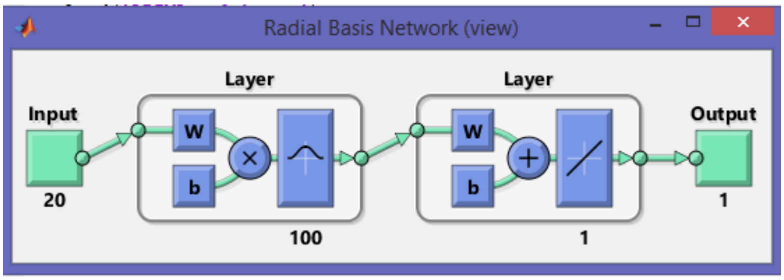


Figure 6.10: RBFNN Architecture for selected attributes

The results show that the performance of the RBFNN model to predict government cash demand in the future increased by utilising the selected attributes. The residual of the proposed RBFNN model for both using initial dataset and selected attributes is presented in Figure 6.11.

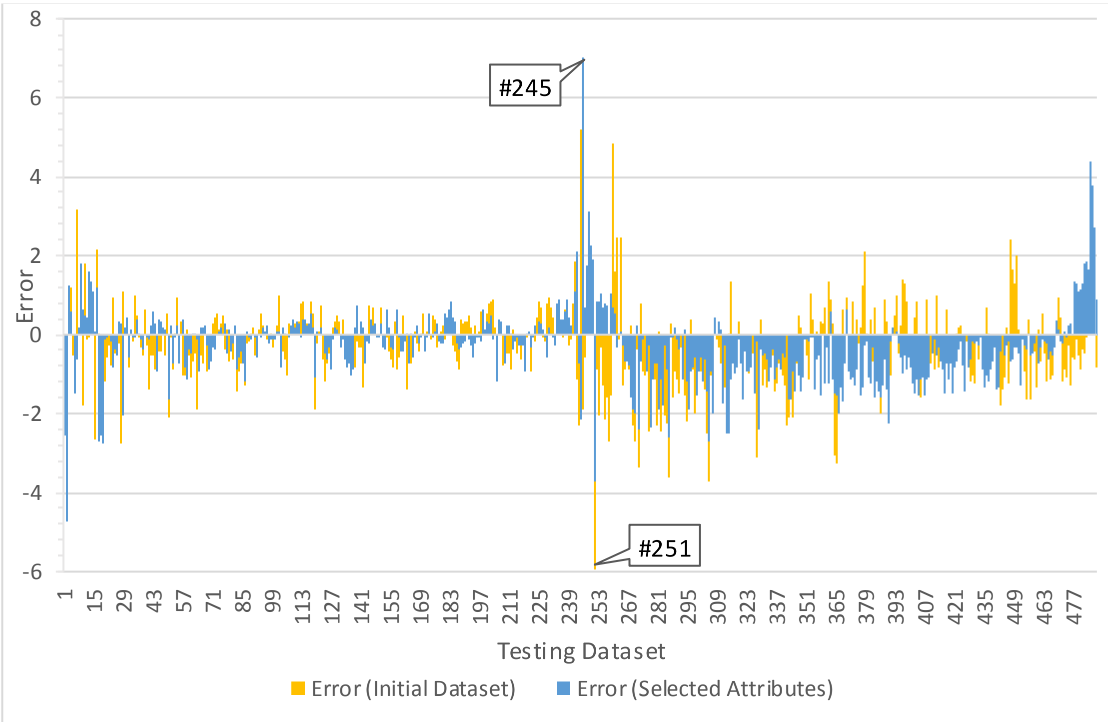


Figure 6.11: The residual of RBFNN model

As detailed in the explanation of Figure 6.2., the description of X-axis and Y-axis of Figure 6.11. are the same as the axes in Figure 6.2. It is obvious from Figure 6.11. that some large deviations happened throughout the testing dataset. The outliers occurred during the first quarter of the budget year (e.g. Observations 2, 245, 251, 260) where the spending units' expenditures were relatively low compared to the other quarters as presented in Figure 6.11. As an illustration, the biggest deviation observed in the testing dataset number 245 was on the 5th of January 2015. The absolute value of the error from both models on the particular day is 5.1869 and 7.0220 for the model that was developed using the initial dataset and selected attributes respectively. It suggests the superiority of the government cash forecasting model that was developed using the initial dataset over the one that was built under the selected attributes. Nevertheless, another obvious figure suggested the contrary. Observation 251 which was on the 13th of January 2015 provides the value of 5.96653 and 3.7076 for the absolute value of the error from the models developed using the initial dataset and selected attributes respectively such that indicates the selected attributes as a better dataset to develop a government cash forecasting model. Both observations demonstrated the alteration of results throughout the testing dataset. Although the errors from both models was varying and overlapping each other along the observation period, further investigation revealed most of the residuals of the RBFNN model developed from the initial dataset were larger than the errors of the model from the selected attributes. The conclusion confirms the result from models' performance evaluations that by utilising the selected attributes data, the RBFNN model is superior compared to the one with the initial dataset.

Furthermore, Figure 6.11. also demonstrates that both models had difficulty in forecasting the expenditure of spending units during the beginning of the budget year where the large deviation between the actual and the forecasted value of the expenditure only happens during the first month of each budget year. The proposition for such a condition can be traced back to the disbursement patterns of the spending units. As explained in Section 3.4., the nature of the type of expenditure influences the spending behaviour of the government agencies in Indonesia. Some expenditure, such as capital expenditure, is embodied in the specific procurement process that may delay the cash disbursement of the spending units up to the end of the budget year. Massive expenditure toward the end of the budget year causes low disbursement at the beginning of the budget year. Policy enhancement on the expenditure procedures might stimulate the spending behaviour such that the spending units' disbursement patterns would be easier to predict. However, the discussion in this regard is outside of the focus of the present study. Despite its limitation, the proposed RBFNN models are capable of producing a reasonable accuracy when measured with the proposed performance measurement evaluation tools.

6.3.2.4. Generalised Regression Neural Network Model

As a member of the radial basis neural network class, GRNN has a radial basis function in its hidden layer. Similarly to RBFNN, GRNN uses a linear activation function in its output layer. However, unlike the RBFNN, the number of neurons in the GRNN's hidden layer is equal to the number of input instances. Therefore, the spread of radial basis functions is the only parameter that the user needs to confirm.

<u>Initial dataset</u>

This study proposed a GRNN network with 0.5 spreads with 1227 neurons as the best model with the initial dataset. The network gave performance scores of 0.648, 0.805, 0.563, 2.084, 2.093, and 1.037 when it was measured with MSE, RMSE, MAE, MAPE, SMAPE, and MASE respectively. GRNN architecture for the initial dataset is shown in Figure 6.12.

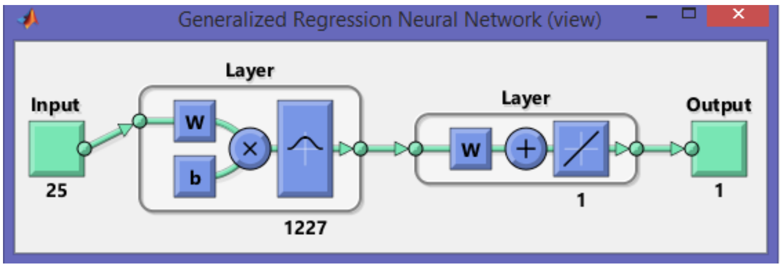


Figure 6.12: GRNN Architecture for initial dataset

<u>Selected Attributes</u>

This study proposed a GRNN network with 0.5 spreads with 1227 neurons as the best model with the selected attributes. The network gave performance scores of 0.522, 0.723, 0.487, 1.832, 1.849, and 0.897 when it was measured with MSE, RMSE, MAE, MAPE, SMAPE, and MASE respectively. GRNN architecture for selected attributes is shown in Figure 6.13.

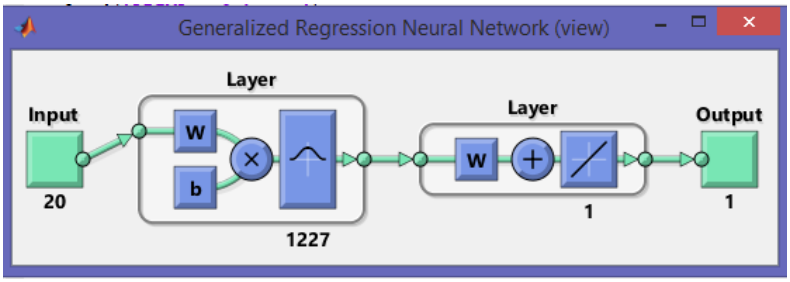


Figure 6.13: GRNN Architecture for initial dataset

The results showed that employing the selected attributes to develop a forecasting model improves the performance of the GRNN model compared to utilising the initial dataset. The residual of the proposed GRNN model for both using initial dataset and selected attributes is presented in Figure 6.14.

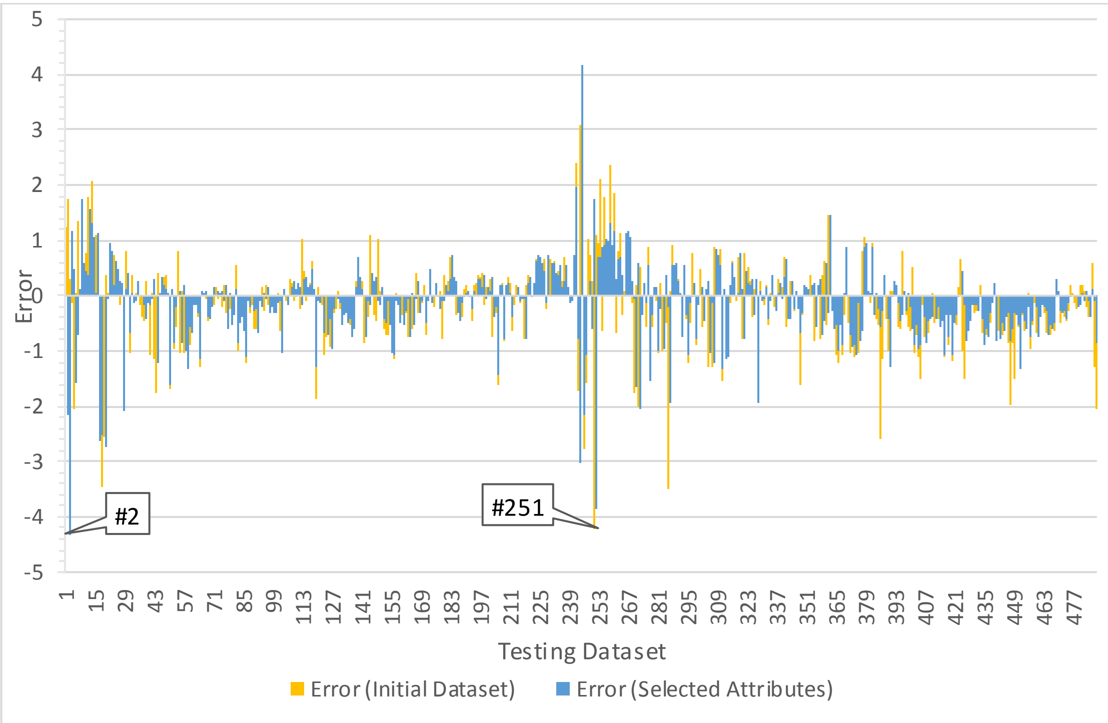


Figure 6.14: The residual of the GRNN model

The legend of Figure 6.14 is the same as the legend of Figure 6.2. It is obvious from Figure 6.14. that some large deviations occurred throughout the testing dataset. The outliers occurred during the first quarter of the budget year (e.g. Observations 2, 18, 251, 244, 245, 251) where the spending units' expenditures were relatively low compared to the other quarters as presented in Figure 6.14. As an illustration, the biggest deviation observed in the testing dataset number 2 was on the 3rd of January 2015. The absolute value of the error from both models on the particular day is 1.7469 and 4.3432 for the model that was developed using the initial dataset and selected attributes respectively. It suggests the superiority of the model that was developed using the initial dataset over the selected attributes. Nevertheless, another obvious figure suggested the contrary. Observation 251 which was on the 13th of January 2015 provides the value of 4.23189 and 3.86306 for the absolute value of the error from the models developed using the initial dataset and selected attributes respectively. It implies the selected attributes as a better dataset to develop a government cash forecasting model. Although both examples indicate the volatility of errors which tend to overlap each other along the observation period, most of the features showed in Figure 6.14 suggest the use of the selected attributes as a better technique to model the future of the spending units' expenditure compared to the initial dataset, when the GRNN method is used to construct a government cash forecasting model. The figure confirms the previous conclusion which is derived from the performance measurements that by utilising the selected attributes data, the GRNN model is superior compared to the one with the initial dataset.

Furthermore, Figure 6.14. also demonstrates that both models had adversity in predicting the expenditure of spending units during the beginning of the budget year where the large deviation between the actual and the forecasted value of the expenditure only occurs during the first month of each budget year. The hypothesis to explain the phenomena can be drawn from the disbursement patterns of the spending units. As explained in Section 3.4., the nature of the type of expenditure influences the spending behaviour of the government agencies in Indonesia. Most spending units' construction projects may be concluded at the end of the budget year due to specific requirements such as procurement procedures that may delay the cash disbursement of the spending units. The spending units' disbursement patterns that is extremely high toward the end of the budget year causes low spending at the beginning of the budget year. Improvement in the procurement process and disbursement mechanism might stimulate the spending behaviour. However, the discussion in this regard is outside of the focus of the present study. Despite its limitation, the proposed GRNN models are capable of producing a reasonable accuracy when they are measured with the proposed performance measurement evaluation tools.

6.3.2.5. Extreme Learning Machine Model

This study took advantage of the work of Huang et al. (2004) in providing ELM's MATLAB codes[3]. ELM is a single-hidden layer feedforward network. Therefore, the only parameter that needs to be set is the number of neurons on its hidden layer.

<u>Initial dataset</u>

This study proposed an ELM network with 50 neurons as the best model with the initial dataset. The network gave performance scores of 1.359, 1.166, 0.828, 3.105, 3.087, and 1.525 when it was measured with MSE, RMSE, MAE, MAPE, SMAPE, and MASE respectively. ELM architecture for the initial dataset is shown in Figure 6.15.

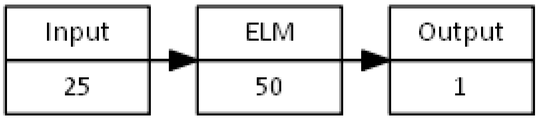


Figure 6.15: The architecture of ELM for the initial dataset

<u>Selected Attributes</u>

This study proposed an ELM network with 95 neurons as the best model with the selected attributes. The network gave performance scores of 0.777, 0.881, 0.656, 2.431, 2.422, and 1.209 when it was measured with MSE, RMSE, MAE, MAPE, SMAPE, and MASE respectively. ELM Architecture for selected attributes is shown in Figure 6.16

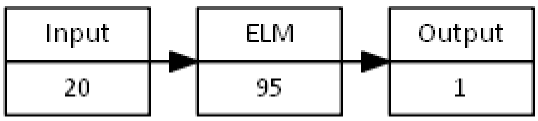


Figure 6.16: The architecture of ELM for the selected attributes

The results show that employing the selected attributes to develop a forecasting model outperforms the model that is built under the initial data when the ELM

[3] Source codes can be found on http://www.ntu.edu.sg/home/egbhuang/elm_codes.html

method is applied. The residual of the proposed ELM model for both using initial dataset and selected attributes is presented in Figure 6.17.

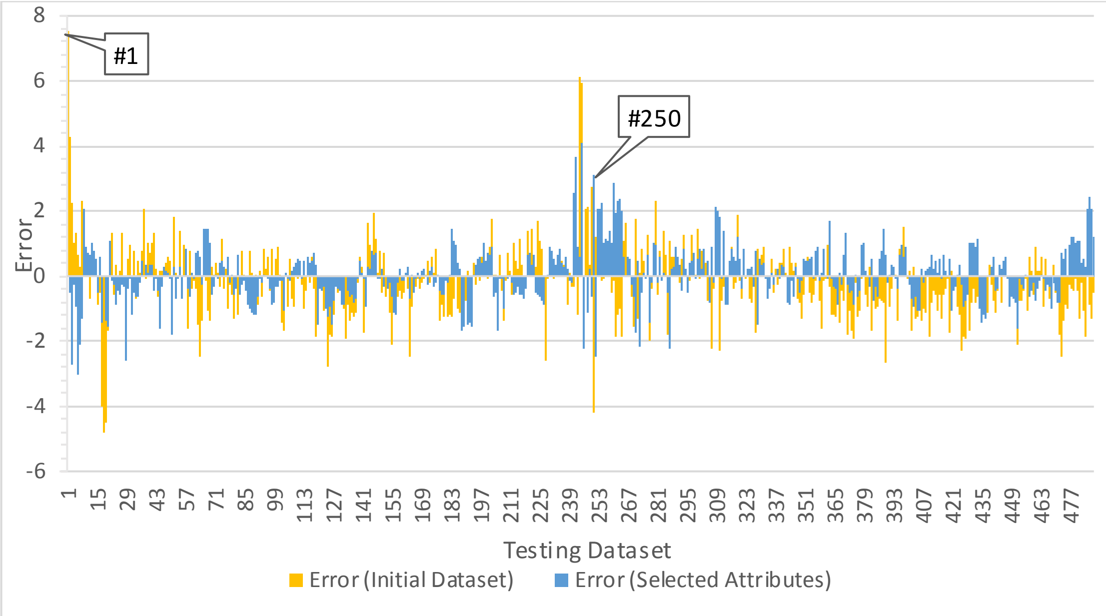


Figure 6.17: The residual of the ELM model

As detailed in the explanation of Figure 6.2., the coordinate setting for Figure 6.17. is identical to the one in Figure 6.2. It is obvious from Figure 6.17. that some large deviations happened throughout the testing dataset. The outliers occurred during the first quarter of the budget year (e.g. Observations 1, 2, 17, 19, 244, 245, 250) where the spending units' expenditures were relatively low compared to the other quarters as presented in Figure 6.17. As an illustration, the biggest deviation observed in the testing dataset number 1 was on the 2nd of January 2014. The absolute value of the error from both models on the particular day is 7.5367 and 0.5034 for the model that was developed using the initial dataset and selected attributes respectively. The most noticeable feature of Figure 6.17 suggests the superiority of the model that was developed using the selected attributes over the one that was built under the initial dataset to project the government cash needed in the future. Nevertheless, another point in the figure suggested the opposite. Observation 250 which was on the 12th of January 2015 provides the value of 2.743282 and 3.126468 for the absolute value of the error from the models developed using the initial dataset and selected attributes respectively. It indicates the initial dataset as a better dataset to develop a government cash forecasting model. Furthermore, both observations illustrated the disparity of results throughout the testing period. The deviations from both models, as shown in Figure 6.17., fluctuate and tend to overlap each other along the observation period. However, more in-depth investigation revealed that most of the errors from the model which developed with the initial dataset are larger than the one with the selected attributes. Therefore, when the ELM technique is employed, the best government cash forecasting model can be attained by using the selected attributes dataset. The figure also confirms the previous conclusion derived from the performance measurements that by utilising the selected attributes data, the ELM model is superior compared to the one with the initial dataset.

Furthermore, Figure 6.17. also demonstrates that both ELM models had an issue to project the expenditure of spending units during the beginning of the budget year where the large deviation between the actual and the forecasted value of the expenditure only occurred during the first month of each budget year. The disbursement patterns of the spending units is more likely to be the cause of the problem. Section 3.4. explained that the spending behaviour of the government agencies in Indonesia is predetermined by the characteristic of the type of expenditure itself. Some expenditure, such as the spending units' construction project, is embodied in the specific procurement process that may delay the cash disbursement of the spending units due to it conclusion at the end of the budget year. Therefore, the accumulated spending toward the end of the budget year causes low disbursement at the beginning of the budget year. In general, the spending behaviour can be catalysed by improving the procurement process and disbursement mechanism. However, the discussion in this regard is outside of the focus of the present study. Despite its limitation, the proposed ELM models are capable of producing a reasonable accuracy when they are measured with the proposed performance measurement evaluation tools.

6.3.2.6. Long Short-Term Memory Model

The present study developed an LSTM network in Python software using the Keras deep learning library. It allows the user to construct an LSTM network with multiple layers (Chollet, 2015). The user needs to set the number of LSTM layers and the number of neurons in each LSTM layer to develop a forecasting model.

Initial dataset

This study proposed an LSTM network with one LSTM layer with 50 neurons as the best model with the initial dataset. The network gave performance scores of 0.502, 0.709, 0.469, 1.765, 1.758, and 0.864 when it was measured with MSE, RMSE, MAE, MAPE, SMAPE, and MASE respectively. LSTM architecture for the initial dataset is shown in Figure 6.18.

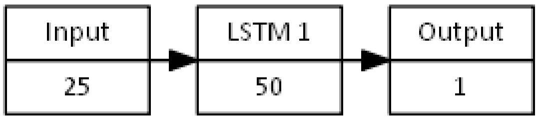


Figure 6.18: The architecture of LSTM for the initial dataset

Selected Attributes

This study proposed an LSTM network with two hidden layers with 50 neurons as the best model with the selected attributes. The network gave performance scores of 0.451, 0.672, 0.468, 1.737, 1.743, and 0.862 when it was measured with MSE, RMSE, MAE, MAPE, SMAPE, and MASE respectively. LSTM architecture for the selected attributes is shown in Figure 6.19.

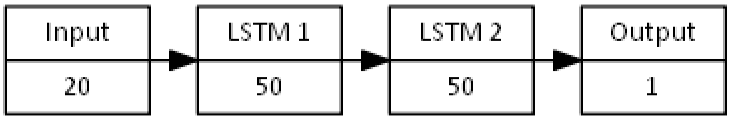


Figure 6.19: The architecture of LSTM for the selected attributes

The results show that the performance of the LSTM model to predict government cash demand in the future increases when the selected attributes are used. The residual of the proposed LSTM model for both using initial dataset and selected attributes is presented in Figure 6.20.

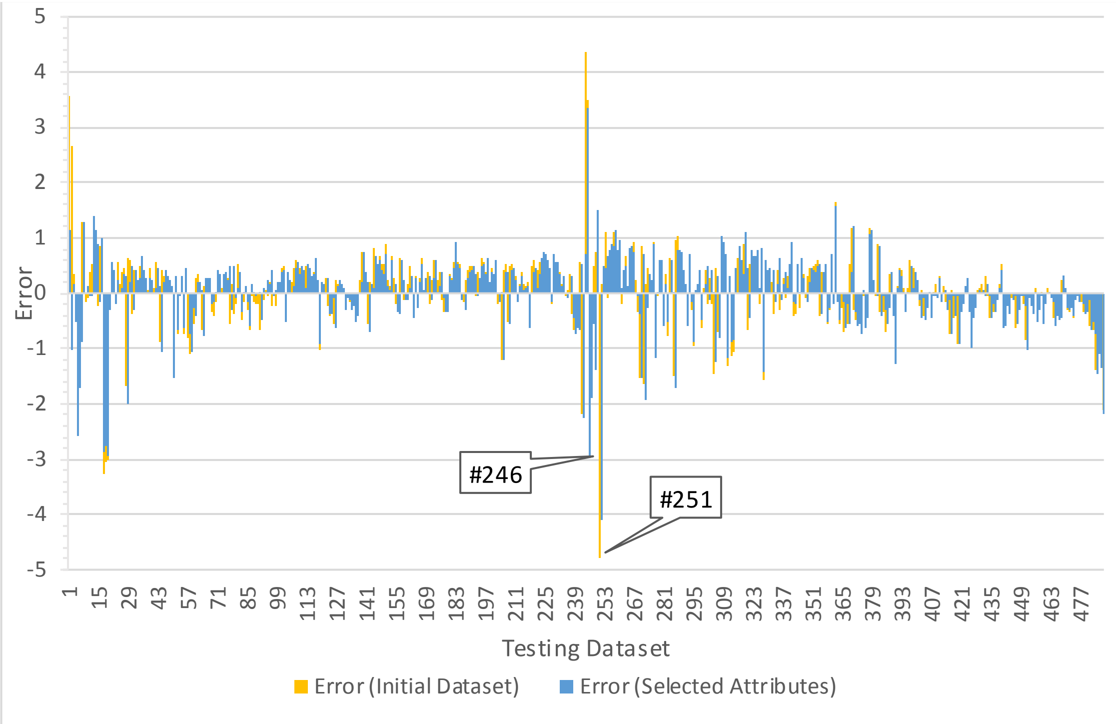


Figure 6.20: The residual of the LSTM model

As detailed in the explanation of Figure 6.2., the X-axis and Y-axis of Figure 6.20. are the same as the axes in Figure 6.2. It is obvious from Figure 6.20. that some large deviations occurred throughout the testing dataset. The outliers occurred during the first quarter of the budget year (e.g. Observations 1, 17, 18, 244, 246, 251) where the spending units' expenditures were relatively low compared to the other quarters as presented in Figure 6.20. As an illustration, the biggest deviation observed in the testing dataset number 251 was on the 13th of January 2015. The absolute value of the error from both models on the particular day is 4.7699 and 4.0834 for the model that was developed using the initial dataset and selected attributes respectively. The most noticeable feature of the figure suggests the superiority of the model that was developed using the selected attributes over the one that was built under the initial dataset to project the government cash needed in the future. Nevertheless, another point in the figure suggested the contrary. The Observation 246 which was on the 6th of January 2015 provides the value of 2.63328 and 2.96768 for the absolute value of the error from the models developed using the initial dataset and selected attributes respectively. It showed that the initial dataset is a better dataset to develop a government cash forecasting model. Moreover, both illustrations demonstrate the difference of results throughout the testing dataset. Although the errors from both models fluctuate and tend to overlap each other along the observation period, Figure 6.20. recommends the used of the selected attributes dataset to build the best government cash forecasting model when the LSTM method is utilised. It is drawn from the fact that most of the errors from the model developed with the initial dataset are larger than the one with the selected attributes. It confirms the previous conclusion which is derived from the performance measurements that by utilising the selected attributes data, the LSTM model is superior compared to the one with the initial dataset.

Furthermore, the large deviation between the actual and the forecasted value of the expenditure only happens during the first month of each budget year in Figure 6.20. also demonstrates that both models had a problem to forecast the expenditure of spending units during the beginning of the budget year. The possible cause of this difficulty can be investigated from the disbursement patterns of the spending units. As explained in Section 3.4., the nature of the type of expenditure influences the spending behaviour of the government agencies in Indonesia. Some capital expenditure, such as spending units' construction projects, may be concluded at the end of the budget yeardue todelay on particular procurement process which affect the cash disbursement of the spending units. The accumulation of spending units' expenditures toward the end of the budget year causes low disbursement at the beginning of the budget year. Policy improvement on the procurement process and disbursement mechanism might stimulate the spending behaviour. However, the discussion in this regard is outside of the focus of the present study. Despite its limitation, the proposed LSTM models are capable of producing a reasonable accuracy when they are measured with the proposed performance measurement evaluation tools.

6.3.2.7. Gated Recurrent Unit Model

This study constructs an LSTM network in Python software using the Keras deep learning library. As described in Section 5.5.2.7., the difference between GRU and LSTM is in the configuration of its layers. Nevertheless, for practical implementation under Keras environment, both techniques are the same. The only parameters that the user needs to set are the number of LSTM layers and the number of neurons in each LSTM layer to develop a forecasting model.

Initial dataset

This study proposed a GRU network with two hidden layers with 50 neurons as the best model with the initial dataset. The network gave performance scores of 0.493, 0.702, 0.452, 1.700, 1.696, and 0.832 when it was measured with MSE, RMSE, MAE, MAPE, SMAPE, and MASE respectively. GRU architecture for the initial dataset is shown in Figure 6.21.

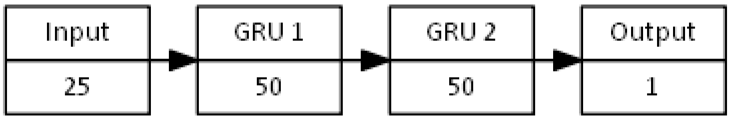


Figure 6.21: The architecture of GRU for the initial dataset

Selected Attributes

This study proposed a GRU network with two hidden layers with 100 neurons as the best model with the selected attributes. The network gave performance scores of 0.424, 0.651, 0.450, 1.671, 1.674, and 0.829 when it was measured with MSE, RMSE, MAE, MAPE, SMAPE, and MASE respectively. GRU architecture for selected attributes is shown in Figure 6.22.

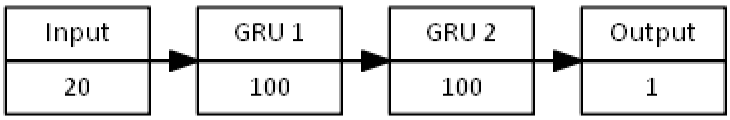


Figure 6.22: The architecture of GRU for the selected attributes

The results show that the performance of the GRU model to predict government cash demand in the future increases by including only the significant attributes. The residual of the proposed GRU model for both using initial dataset and selected attributes is presented in Figure 6.23.

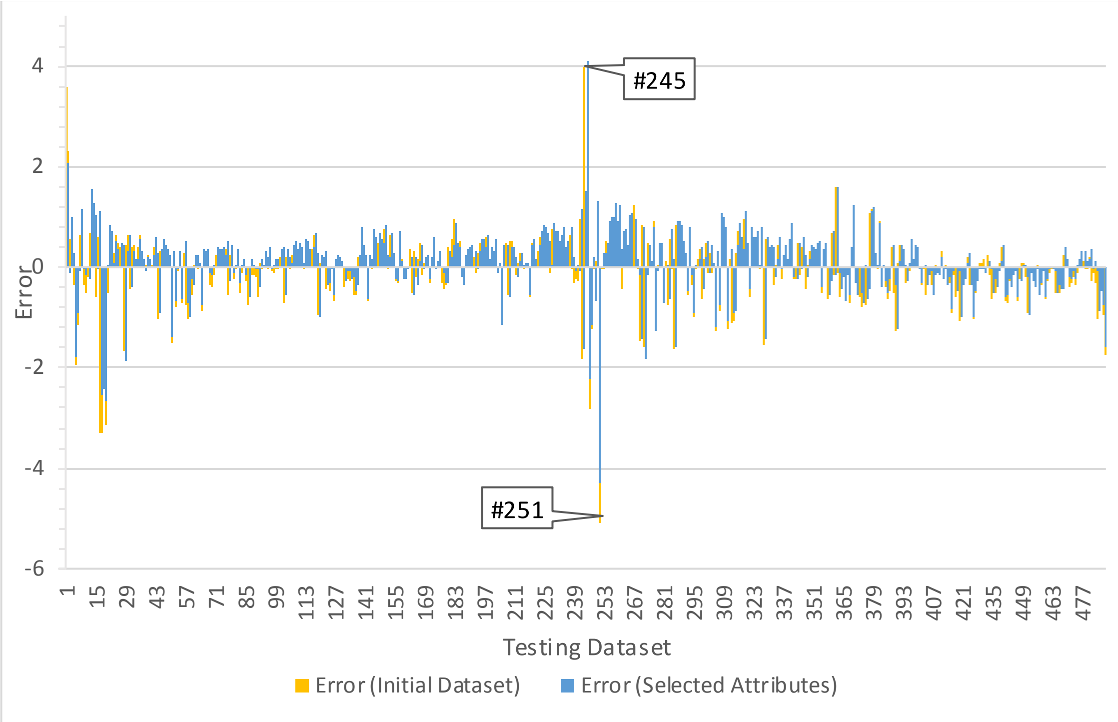


Figure 6.23: The residual of the GRU model

The coordinate arrangement for Figure 6.23. is the same as the axes in Figure 6.2. It is obvious from Figure 6.23. that some large deviations happened throughout the testing dataset. The outliers occurred during the first quarter of the budget year (e.g. Observations 1, 17, 18, 244, 245, 251) where the spending units' expenditures were relatively low compared to the other quarters as presented in Figure 6.23. As an illustration, the biggest deviation observed in the testing dataset number 251 was on the 13th of January 2015. The absolute value of the error from both models on the particular day is 5.0716 and 4.2760 for the model that was developed using the initial dataset and selected attributes respectively. The most noticeable feature of the figure suggests the superiority of the model that was developed using the selected attributes over the one that was built under the initial dataset to project the government cash needed in the future. Nevertheless, another point in the figure suggested the contrary. Observation 245 which was on the 5th of January 2015 provides the value of 3.729708 and 4.086908 for the absolute value of the error from the models developed using the initial dataset and selected attributes respectively such that compared to the selected attributes, the initial dataset is a better dataset to develop a government cash forecasting model. Both observations illustrated the variation of results throughout the testing dataset. Although the deviations from both models oscillate and overlap each other along the observation period, Figure 6.23. supports the suggestion to use the selected attributes dataset in constructing the best government cash forecasting model under the GRU technique since most of the errors from the initial dataset model are larger than the errors of the model from the selected attributes. It confirms the previous conclusion derived from the performance measurements that by utilising the selected attributes data, the GRU model is superior compared to the one with the initial dataset.

Furthermore, Figure 6.23. shows that disregarding the data used, forecasting the expenditure of spending units during the beginning of the budget year is challenging for both models. It can be seen from the large deviation between the actual and the forecasted value of the expenditure that only happens during the first month of each budget year. The disbursement patterns of the spending units offer a proposition on possible cause of the phenomena. As explained in Section 3.4., the characteristic of the type of expenditure influences the spending behaviour of the government agencies in Indonesia. Some expenditures, such as capital expenditure, are embodied in the specific procurement process that may delay the cash disbursement of the spending units due to the completion of the spending units' construction project. As a consequence, the majority of spending units' expenditures are accumulated toward the end of the budget year which causes low disbursement at the beginning of the budget year. Improvement of the expenditure procedures might stimulate the spending behaviour. However, the discussion in this regard is outside of the focus of the present study. Despite its limitation, the proposed GRU models are capable of producing a reasonable accuracy when they are measured with the proposed performance measurement evaluation tools.

6.3.3 Hybrid Model

As described in Section 5.5.3., the hybrid model lays its theory on the assumption that every time series data are embodied with both linear and nonlinear components, which shape its patterns. The relationship between both elements may follow additive and multiplicative approaches. However, there is no justification for defining whether the linear and nonlinear relationship of a time series data follows an additive or multiplicative approach. Dealing with these approaches, Zhang (2003) proposes a hybrid model based on the additive approach while Wang et al. (2013) employ the multiplicative model. Therefore, this study utilises both approaches to investigate the best hybrid model to predict the future of government cash requirements.

The hybrid model combines more than one method to construct a forecasting model. Specifically, for this study, it is coalescing ARIMA model and ANN technique. In practice, the hybrid model consists of two parts. Firstly, the data is used to develop a forecasting model following an ARIMA method. Once a model is established, the residuals of the ARIMA model are set as input and target to the ANN technique in the next step.

In this study, the hybrid model uses the ARIMAX model for the first phase of hybrid modelling. Once the forecasted values are estimated, the residuals of the ARIMAX model are generated based on the selected approach for the hybrid model. The residuals are calculated based on equation 5.54 for the additive model and equation 5.57 for the multiplicative model. Since the assumption is that for every time series data consist of the linear and nonlinear components, the residual of the linear model is the nonlinear components. In the additive model, a time series data can be estimated as a summation of linear and nonlinear components. Therefore, the residual of the linear model that consists of the nonlinear components can be calculated by subtracting the actual values of the time series data with its forecasted values from the ARIMAX model. On the other hand, the multiplicate model considers a time series data as a multiplication of linear and nonlinear components. Hence, the residual of the linear model can be calculated by dividing the actual values of the time series data with its forecasted values from the ARIMAX model. Once the set of the residuals is assembled, it is fed into the next step which is modelling using the ANN technique. For modelling using the ANN technique, there is only one variable to be treated as input and output that is the residual of the ARIMAX model. Therefore, NARNN is chosen as the modelling method in the hybrid model due to its ability to train a network and make a prediction from its past values.

The first part of the hybrid model, which is modelling using the ARIMAX model, is as reported in Section 6.3.1. This section only discusses the later stage of the hybrid model that includes the development of a hybrid model using NARNN, and the performance evaluation of the hybrid model. The NARNN was developed under the MATLAB. With the help of the MATLAB toolbox, the number of hidden layers, the number of neurons in each hidden layer, and the number of feedback delays are the parameters that the user needs to set.

<u>**Initial dataset**</u>

For the additive model, the NARNN technique proposed a network with four hidden layers with five neurons in each layer and four lags of feedback delays as the best model with the initial dataset as presented in Figure 6.24. Overall, the hybrid model with the additive model gave performance scores of 1.239, 1.113, 0.648, 2.484, 2.625, and 1.194 when it was measured with MSE, RMSE, MAE, MAPE, SMAPE, and MASE respectively.

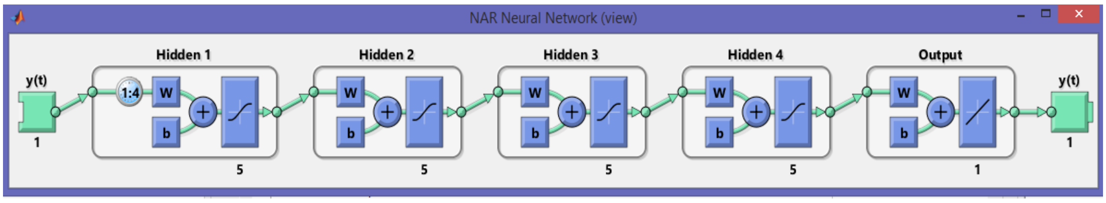


Figure 6.24: NARNN architecture for the initial dataset (Additive model)

Furthermore, the multiplicative model suggested the NARNN network with two hidden layers, five neurons in each hidden layer, and four lags of feedback delays. Figure 6.25. displays the topology of NARNN for the initial dataset based on the multiplicative model. The multiplicative hybrid model gave performance scores of 3.009, 1.735, 1.452, 5.406, 5.401, and 2.676 when it was measured with MSE, RMSE, MAE, MAPE, SMAPE, and MASE respectively.

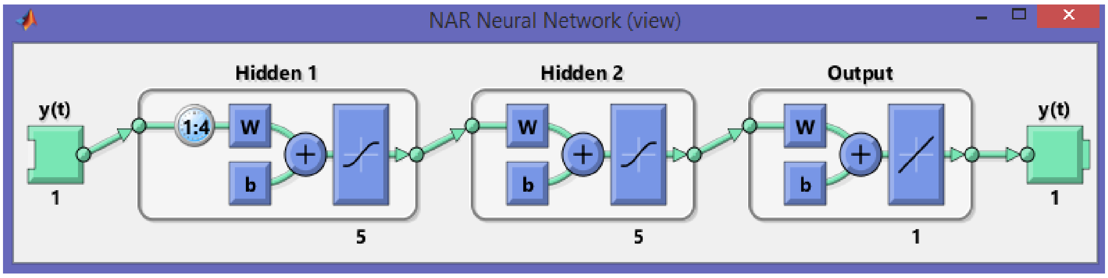


Figure 6.25: NARNN architecture for the initial dataset (Multiplicative model)

Performance of the models suggests that the relationship between linear and nonlinear components of the initial dataset follows the additive model. It can be seen in Table 6.6. where the performance of the hybrid model that was built using the initial dataset and based on the additive model was better than the one that was built on the multiplicative model. Therefore, for the initial dataset, this study proposes the hybrid model that follows the additive model as the best approach to model the government cash forecasting.

Table 6.6: Performance comparison of additive and multiplicative models (the initial dataset)

Model	MSE	RMSE	MAE	MAPE	SMAPE	MASE
Additive	**1.239**	**1.113**	**0.648**	**2.484**	**2.625**	**1.194**
Multiplicative	3.009	1.735	1.452	5.406	5.401	2.676

Selected Attributes

Similar to the initial dataset, this study also investigates the best hybrid model using the selected attributes by comparing the performance of both models that are developed based on additive and multiplicative models.

For the additive model, the best topology of NARNN for the selected attributes proposed a network with three hidden layers with five neurons in each layer and two lags of feedback delays as the best model with the selected attributes. The best government cash forecasting model developed under the hybrid model gave performance scores of 1.159, 1.076, 0.582, 2.268, 2.421, and 1.072 when it was measured with MSE, RMSE, MAE, MAPE, SMAPE, and MASE respectively. The architecture of NARNN which is built on the additive model for the selected attributes is as shown in Figure 6.26.

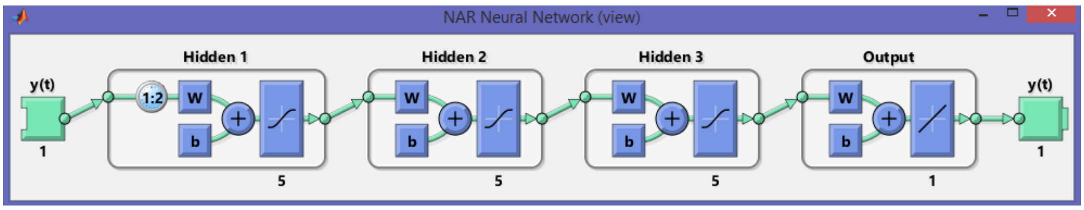


Figure 6.26: NARNN architecture for the selected attributes (additive model)

Furthermore, the multiplicative model suggested the NARNN network with three hidden layers, five neurons in each hidden layer, and five lags of feedback delays. Figure 6.27. displays the topology of NARNN for the selected attributes based on the multiplicative model. The multiplicative hybrid model gave performance scores of 1.468, 1.211, 0.718, 2.783, 2.943, and 1.323 when it was measured with MSE, RMSE, MAE, MAPE, SMAPE, and MASE respectively.

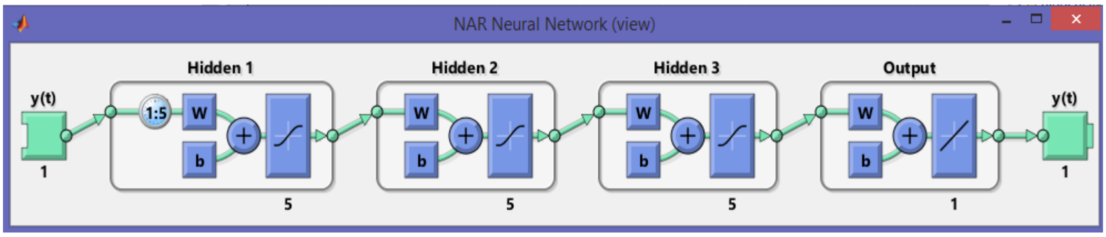


Figure 6.27: NARNN architecture for the selected attributes (Multiplicative model)

Performance of the models suggests that the relationship between the linear and nonlinear components of the selected attributes follows the additive model. Table 6.7. displays the performance of the hybrid model that was built using the selected attributes and concludes that using the additive model is a better technique compared to the one that is developed under the multiplicative approach. Therefore, for the selected attributes, this study proposed the hybrid model that follows the additive model as the best approach to model the government cash forecasting.

Table 6.7: Performance comparison of additive and multiplicative models (the selected attributes)

Model	MSE	RMSE	MAE	MAPE	SMAPE	MASE
Additive	**1.159**	**1.076**	**0.582**	**2.268**	**2.421**	**1.072**
Multiplicative	1.468	1.211	0.718	2.783	2.943	1.323

Overall, the experiments suggest the use of the additive hybrid model to construct a government cash forecasting model for both the initial dataset and the selected attributes. Table 6.6. and 6.7. show that utilising only the significant attributes in developing a model to predict the future of government cash demand is superior compared to the one using the initial data. The residual of the proposed hybrid model for both using initial dataset and selected attributes is presented in Figure 6.28.

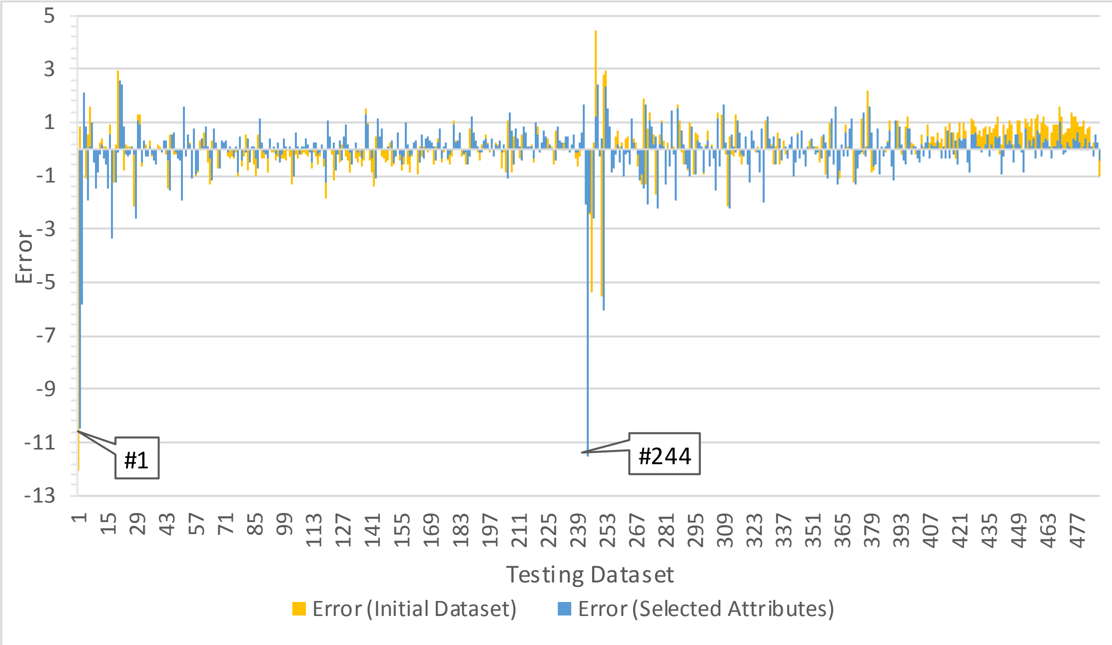


Figure 6.28: The residual of the Hybrid model

As detailed in the explanation of Figure 6.2., the X-axis and Y-axis of Figure 6.28. are the same as the axes in Figure 6.2. It is obvious from Figure 6.28. that some large deviations occurred throughout the testing dataset. The outliers occurred during the first quarter of the budget year (e.g. Observations 1, 2, 244, 248, 251) where the spending units' expenditures were relatively low compared to the other quarters as presented in Figure 6.28. As an illustration, the biggest deviation observed in the testing dataset number 1 was on the 2nd of January 2014. The absolute value of the error from both models on the particular day is 12.0171 and 10.4358 for the model that was developed using the initial dataset and selected attributes respectively. The most noticeable feature of the figure suggests the superiority of the model that was developed using the selected attributes over the one that was built under the initial dataset to project the government cash needed in the future. Nevertheless, another point in the figure suggested the contrary. Observation 244 which was on the 1st of January 2015 provides the value of 10.0703 and 11.5559 for the absolute value of the error from the models developed using the initial dataset and selected attributes respectively. It indicates the initial dataset as a better dataset to develop a government cash forecasting model. Both observations represented the variation of results throughout the testing dataset. While the errors from both models fluctuate and tend to overlap each other along the observation period, Figure 6.28. shows that the best government cash forecasting model utilising the hybrid model can be achieved by using the selected attributes. It is observable from the fact that most of the errors of the model constructed with the initial dataset are larger than the one with the selected attributes. It confirms the previous conclusion derived from the performance measurements of the models described earlier, that the hybrid model performs better when utilising the selected attributes data compared to the one with the initial dataset.

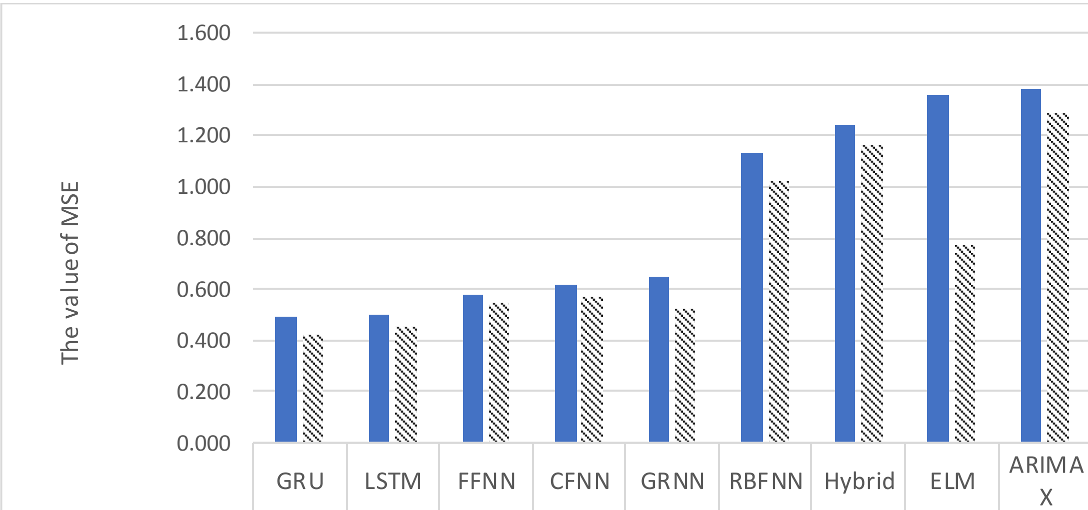
Furthermore, Figure 6.28. displays that both models had difficulty in forecasting the expenditure of spending units during the beginning of the budget year where the large deviation between the actual and the forecasted value of the expenditure only happens during the first month of each budget year. The proposition for such a condition can be traced back to the disbursement patterns of the spending units. As explained in Section 3.4., the nature of the type of expenditure influences the spending behaviour of the government agencies in Indonesia. Some capital expenditure, such as the spending units' construction projects, embodied in the specific procurement process that may be concluded at the end of the budget year. It might delay the cash disbursement of the spending units up to the last month of the budget year. This pattern make low disbursement at the beginning of the budget year. Improvement on the procurement process and disbursement mechanism might stimulate the spending behaviour. However, the discussion in this regard is outside of the focus of the present study. Despite their limitation, the proposed hybrid models are capable of producing a reasonable accuracy when they are measured with the proposed performance measurement evaluation tools.

6.4. Performance Evaluation

Once all proposed methods are estimated individually, the next step is to examine the best model of government cash forecasting by comparing the performance of each model based on performance evaluation measurements as described in Section 5.5. Six performance measurements are used in this study as the representation of each approach proposed by Hyndman and Athanasopoulos (2018), namely (1) scale-dependent errors, (2) percentage errors, and (3) scaled errors. The following section is dedicated to reporting the comparison of the models' performances for each performance evaluation measurement respectively hence the best procedure to build a government cash forecasting model with the accuracy that meets an acceptable level of materiality for the cash manager to be identified.

6.4.1 Mean Squared Error

Figure 6.29. shows the performance of all proposed models measured by MSE. Overall, the lowest score was achieved when the forecasting model was built using the GRU method on the selected attributes at 0.424, while a model that was constructed using the ARIMAX technique on the initial dataset turned out to be the worst model at 1.380. For the initial dataset, the methods that are used in the present study listed from the lowest to the highest value of MSE score are GRU, LSTM, FFNN, CFNN, GRNN, RBFNN, hybrid model, ELM, and ARIMAX, while for the selected attributes they are GRU, LSTM, GRNN, FFNN, CFNN, ELM, RBFNN, hybrid model, and ARIMAX models. Moreover, all models built on the selected attributes are superior to the one that was developed using the initial dataset when it is measured individually.



	GRU	LSTM	FFNN	CFNN	GRNN	RBFNN	Hybrid	ELM	ARIMAX
Initial Dataset	0.493	0.502	0.581	0.618	0.648	1.130	1.239	1.359	1.380
Selected Attributes	0.424	0.451	0.544	0.567	0.522	1.025	1.159	0.777	1.290

Figure 6.29: MSE performance scores

The lowest value of MSE means the best model. Figure 6.29. confirms the use of the GRU model on the selected attributes as the best procedure to develop a government cash forecasting model with accuracy that meets an acceptable level of materiality for the cash manager based on the MSE measurement. The graph also provides a visual evidential that deep learning class of neural network techniques (e.g. LSTM and GRU) performed better than the shallow class (e.g. FFNN, CFNN, RBFNN, and GRNN). Interestingly, the results reveal that the hybrid model failed to be the best model for both the initial dataset and the selected attributes. The hybrid model only performs better than the ELM and ARIMAX method for the initial dataset and ARIMAX method for the selected attributes.

The MSE measurement on the hybrid model supports the finding of Taskaya-Temizel and Casey (2005). The research argued that contrary to its fame, combining several methods into a hybrid model is not always superior to the single method. The rationale for this fact lies on the underlying component of the time series data. As described in Chapter 5., each method has its own strength in capturing the linear or nonlinear patterns of the time series data. The ARIMA based method is well known for its competence in apprehending the linear element of the time series data, while the ANN based technique is capable of handling the non-linear patterns. In both cases linear or nonlinear patterns exist in a time series data, the hybrid method is the best procedure to build a forecasting model. The MSE measurement result suggests the best model is the one that is developed using the ANN method. The fact implies that, based on the MSE measurement, the disbursement patterns in Indonesia tend to be embodied with the nonlinear component. However, extended discussion on this subject is outside this study's scope.

6.4.2 Root Mean Squared Error

The RMSE scores for all proposed models are presented in Figure 6.30. Based on the RMSE scores, developing a forecasting model which follows the GRU method on the selected attributes provides the lowest value at 0.651 and a model that is constructed using the ARIMAX technique on the initial dataset gives the highest value at 1.175. The same as the MSE metric, GRU, LSTM, FFNN, CFNN, GRNN, RBFNN, Hybrid model, ELM, and ARIMAX are the initial dataset-based models with the lowest to the highest value of RMSE. While for the selected attributes, GRU, LSTM, GRNN, FFNN, CFNN, ELM, RBFNN, hybrid model, and ARIMAX models are named the lowest to the highest value of RMSE. Moreover, all models built on the selected attributes are superior to the one that was developed using the initial dataset when it is measured individually.

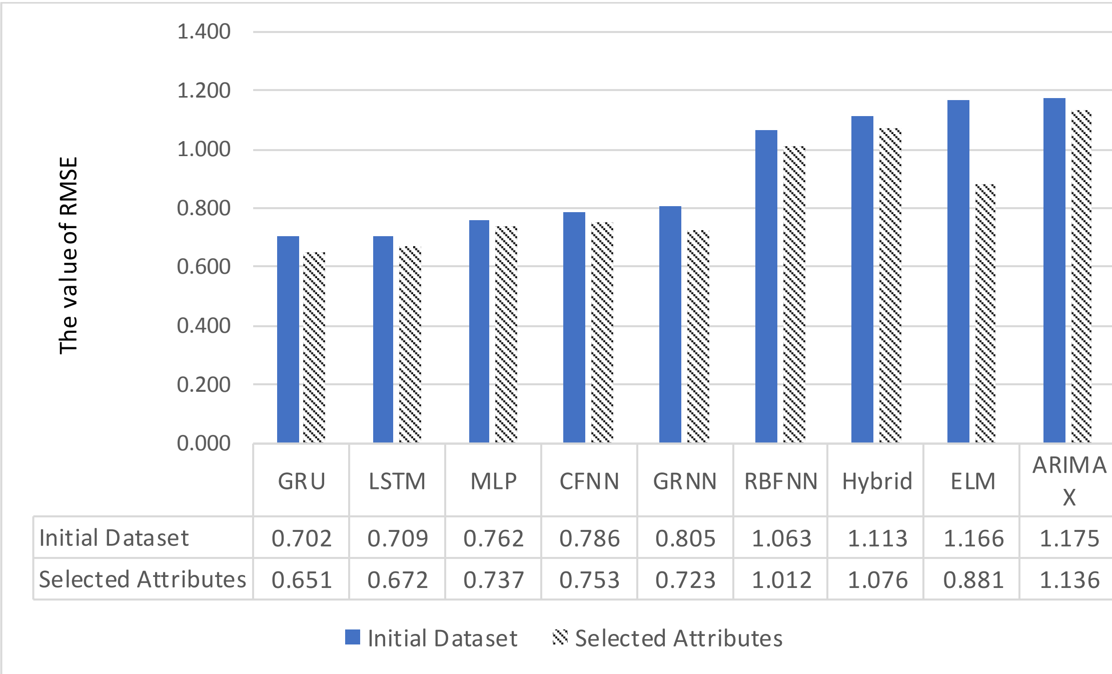


Figure 6.30: RMSE performance scores

In RMSE, the best model is denoted from its lowest value. Figure 6.30. supports the use of the GRU model on the selected attributes as the best procedure to develop a government cash forecasting model with accuracy that meets an acceptable level of materiality for the cash manager based on the RMSE measurement. Figure 6.30. also provides evidence to the superiority of the deep learning class over the shallow class neural network exhibited based on RMSE measurement. Parallel with the MSE measurement, the hybrid model only performs better than the ELM and ARIMAX method for the initial dataset and ARIMAX method for the selected attributes.

The RMSE measurement on the hybrid model supports the finding of Taskaya-Temizel and Casey (2005). The research argued that contrary to its fame, combining several methods into a hybrid model is not always superior to the single method. The rationale for this fact lies in the underlying component of the time series data. As described in Chapter 5., each method has its own strength in capturing the linear or nonlinear patterns of the time series data. When the linear element is prominent in a time series data, the ARIMA-based method is the best technique to use. On the other hand, the ANN-based technique is suitable in handling a time series data with the non-linear patterns. Nevertheless, in the case of both linear or nonlinear patterns existing in a time series data, the hybrid method is the best procedure to build a forecasting model. The RMSE measurement result suggests the best model is the one that is developed using the ANN method. The fact implies that, based on the RMSE measurement, the disbursement patterns in Indonesia tend to be embodied with the nonlinear component. However, an extended discussion on this subject is outside this study's scope.

6.4.3 Mean Absolute Error

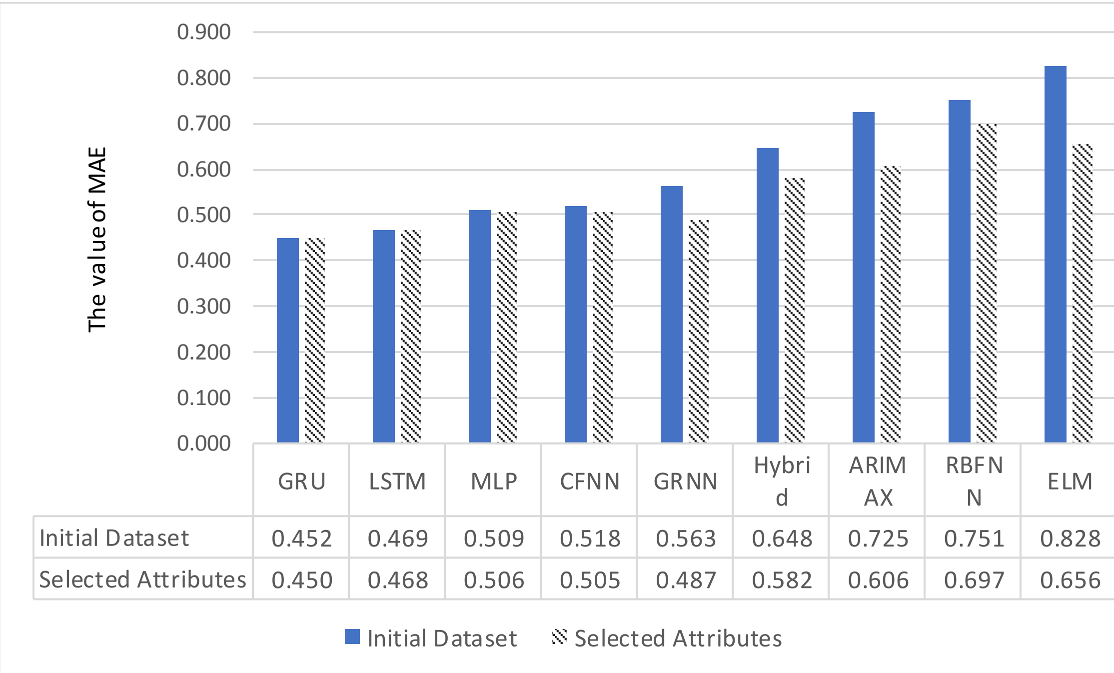


Figure 6.31: MAE performance scores

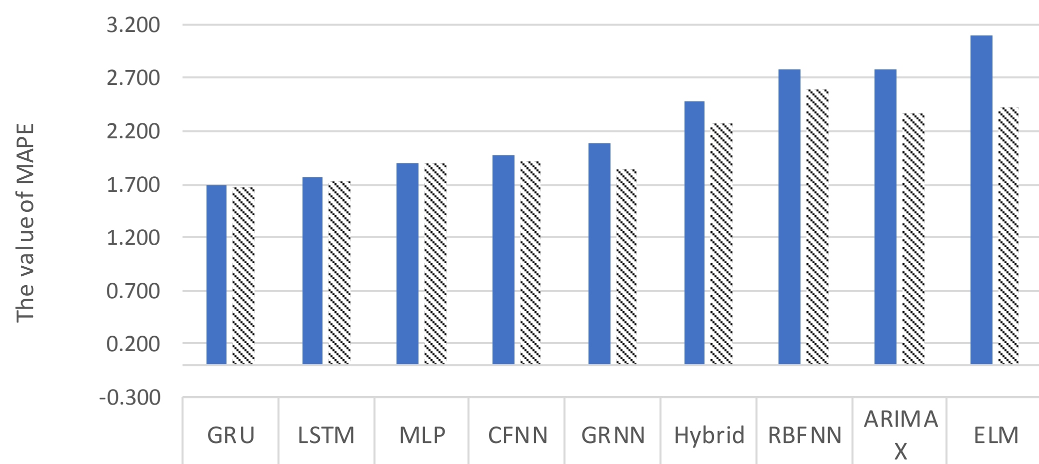
The performances of all proposed models measured by MAE are presented in Figure 6.31. Identical to the previous measurement tools, applying the GRU method on the selected attributes is proven to have the ability to produce the lowest score of the MAE at 0.450. On the other hand, the highest score of MAE at 0.828 is achieved by a model that was developed using the ELM technique on the initial dataset. For the initial dataset, the methods that are used in the present study listed from the lowest to the highest value of MAE score are GRU, LSTM, FFNN, CFNN, GRNN, Hybrid model, ARIMAX, RBFNN, and ELM, whilst for the selected attributes they are GRU, LSTM, GRNN, CFNN, FFNN, Hybrid model, ARIMAX, ELM, and RBFNN models. Consistent with the preceding measurements, employing the selected attributes to construct a government cash forecasting model was confirmed to be a better approach compared to using the initial dataset on the same proposed method.

The lowest value of MAE represents the best model. Figure 6.31. suggests the use of the GRU model on the selected attributes as the best procedure to develop a government cash forecasting model with accuracy that meets an acceptable level of materiality for the cash manager based on the MAE measurement. Figure 6.31. affirmed, based on the MAE measurements, the dominance of the deep learning class over the shallow learning networks which is identical to prior findings. Although utilising the hybrid technique on both the initial dataset and the selected attributes is prominent to increase the performance of the model compared to ARIMAX, RBFNN, and ELM methods, it fails to be the best model to predict the future government cash requirements.

In contrast to its popularity, Taskaya-Temizel and Casey (2005) reported that joining several methods into a hybrid model is not always superior to the single method. The MAE measurement on the hybrid model supports the finding of Taskaya-Temizel and Casey (2005). The proposed hybrid models are unable to provide a model with the best accuracy. The justification for this circumstance can be proposed from the underlying component of the time series data. As described in Chapter 5., each method has its own strength in capturing the linear or nonlinear patterns of the time series data. The ARIMA-based method is well known for its competence in apprehending the linear element of the time series data, while the ANN-based technique is capable of handling the non-linear patterns. In the case of both linear or nonlinear patterns existing in a time series data, the hybrid method is the best procedure to build a forecasting model. The MAE measurement result suggests the best model is the one that is developed using the ANN method. The fact implies that, based on the MAE measurement, the disbursement patterns in Indonesia tend to be embodied with the nonlinear component. However, an extended discussion on this subject is outside this study's scope.

6.4.4 Mean Absolute Percentage Error

Figure 6.32. provides the MAPE score of all proposed models. The MAPE measures the GRU method on the selected attributes as the lowest value of MAPE at 1.671 while the model that was developed using the ELM technique on the initial dataset gave the highest score at 3.105. Moreover, GRU, LSTM, FFNN, CFNN, GRNN, Hybrid model, RBFNN, ARIMAX, and ELM models are the initial dataset-based models with the lowest to the highest value of MAPE. While for the selected attributes, GRU, LSTM, GRNN, FFNN, CFNN, Hybrid model, ARIMAX, ELM, and RBFNN models are named the lowest to the highest value of RMSE. In addition, in agreement with the earlier described measurement criteria, when the selected attributes are employed to create a forecasting model, its performance is increased compared to the one that used the initial dataset.



	GRU	LSTM	MLP	CFNN	GRNN	Hybrid	RBFNN	ARIMAX	ELM
Initial Dataset	1.700	1.765	1.900	1.979	2.084	2.484	2.781	2.786	3.105
Selected Attributes	1.671	1.737	1.893	1.915	1.832	2.268	2.591	2.366	2.431

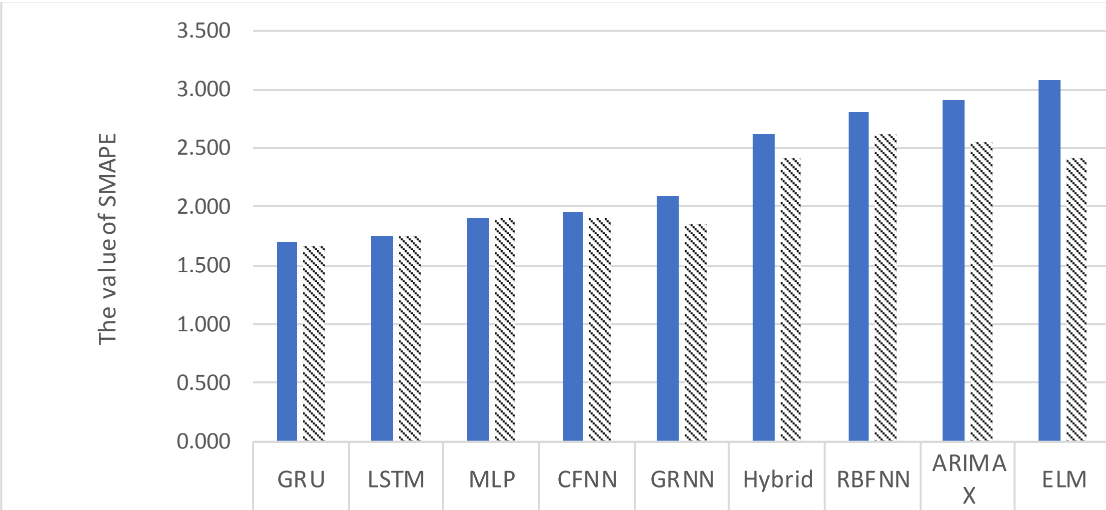
Figure 6.32: MAPE performance scores

The lowest value of MAPE represents the best model. Figure 6.32. confirms the use of the GRU model on the selected attributes as the best procedure to develop a government cash forecasting model with accuracy that meets an acceptable level of materiality for the cash manager based on the MAPE measurement. The graph also provides evidence that the deep learning network performed better than the shallow learning networks when measured by the MAPE score. The performance of the hybrid model is better than RBFNN, ARIMAX, and ELM models for both the initial data and the selected attributes. However, the hybrid model fails to be the best forecasting model.

Taskaya-Temizel and Casey (2005) argued that contrary to its fame, combining several methods into a hybrid model is not always superior to the single method. The result from the MAPE measurement supports the finding where the hybrid method fails to produce the best forecasting model compared to other proposed methods. The rationale for this result can be investigated from the underlying component of the time series data. As described in Chapter 5., each method has its own strength in capturing the linear or nonlinear patterns of the time series data. The ARIMA based method is well known for its competence in apprehending the linear element of the time series data, while the ANN-based technique is capable of handling the non-linear patterns. In the case of both linear or nonlinear patterns existing in a time series data, the hybrid method is the best procedure to build a forecasting model. The MAPE measurement result suggests the best model is the one that is developed using the ANN method. The fact implies that, based on the MAPE measurement, the disbursement patterns in Indonesia tend to be embodied with the nonlinear component. However, an extended discussion on this subject is outside this study's scope.

6.4.5 Symmetric Mean Absolute Percentage Error

The performances of all proposed models measured by SMAPE is presented in Figure 6.33. The model that utilised the GRU method on the selected attributes provides the lowest value of the SMAPE at 1.674. The model that was developed using the ELM technique on the initial dataset gave the highest score of SMAPE at 3.087. Moreover, for the initial dataset, the methods that are used in the present study listed from the lowest to the highest value of SMAPE score are GRU, LSTM, FFNN, CFNN, GRNN, Hybrid model, RBFNN, ARIMAX, and ELM models, while for the selected attributes they are GRU, LSTM, GRNN, FFNN, CFNN, Hybrid model, ELM, ARIMAX, and RBFNN models. Furthermore, agreeing with the earlier described measurement criteria, when the selected attributes are employed to create a forecasting model, its performance is increased compared to the initial dataset.



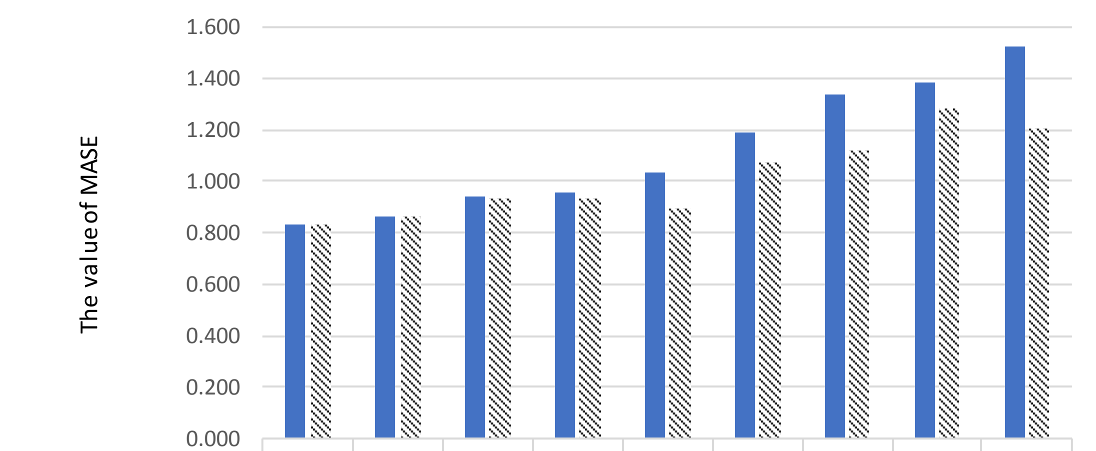
	GRU	LSTM	MLP	CFNN	GRNN	Hybrid	RBFNN	ARIMAX	ELM
Initial Dataset	1.696	1.758	1.911	1.960	2.093	2.625	2.811	2.915	3.087
Selected Attributes	1.674	1.743	1.909	1.910	1.849	2.421	2.612	2.548	2.422

Figure 6.33: SMAPE performance scores

Based on the SMAPE, the best model is denoted by the lowest value. Figure 6.33. confirms the use of the GRU model on the selected attributes as the best procedure to develop a government cash forecasting model with accuracy that meets an acceptable level of materiality for the cash manager based on the SMAPE measurement. The graph also provides evidence that the deep learning network performed better than the shallow learning networks when measured by the SMAPE score. Like the other metrics, the model that was built on the hybrid approach, for both the initial dataset and the selected attributes, failed to be the best model and only performed better against the models that were constructed using the RBFNN, ARIMAX, and ELM method.

Identical to the previous results, the failure of the hybrid model to become the best method in forecasting the government cash requirement when measured with the SMAPE supports the finding of Taskaya-Temizel and Casey (2005). The research argued that contrary to its fame, combining several methods into a hybrid model is not always superior to the single method. The argument for this result can be analysed from the underlying component of the time series data. As described in Chapter 5., each method has its own capacity in estimating the linear or nonlinear patterns of a time series data. The ARIMA based method is well known for its competence in apprehending the linear element of the time series data, while the ANN-based technique is capable of handling the non-linear patterns. In the case of both linear or nonlinear patterns existing in a time series data, the hybrid method is the best procedure to build a forecasting model. The SMAPE measurement result suggests the best model is the one that is developed using the ANN method. Therefore, based on the SMAPE measurement, the disbursement patterns in Indonesia tend to be embodied with the nonlinear component. However, an extended discussion on this subject is outside this study's scope.

6.4.6 Mean Absolute Scaled Error



	GRU	LSTM	MLP	CFNN	GRNN	Hybrid	ARIMAX	RBFNN	ELM
Initial Dataset	0.832	0.864	0.938	0.954	1.037	1.194	1.336	1.385	1.525
Selected Attributes	0.829	0.862	0.933	0.931	0.897	1.072	1.117	1.285	1.209

Initial Dataset   Selected Attributes

Figure 6.34: MASE performance scores

Figure 6.34. shows the performances of all proposed models measured by MASE. The MASE score proposes that the lowest value of the MASE at 0.829 is achieved by the GRU model on the selected attributes. It also confirms that the ELM model on the initial dataset gives the highest score of MASE at 1.525. For the initial dataset, the

methods that are used in the present study listed from the lowest to the highest value of MASE score are GRU, LSTM, FFNN, CFNN, GRNN, Hybrid model, ARIMAX, RBFNN, and ELM models, while for the selected attributes they are GRU, LSTM, GRNN, CFNN, FFNN, Hybrid model, ARIMAX, ELM, and RBFNN models. Furthermore, in line with the earlier described measurement criteria, utilising the selected attributes to create a forecasting model increases its performance compared to the model that used the initial dataset.

In MASE, the best model is denoted from its lowest value. Figure 6.34. supports the use of the GRU model on the selected attributes as the best procedure to develop a government cash forecasting model with accuracy that meets an acceptable level of materiality for the cash manager based on the MASE measurement. Equivalent to the other proposed measurements, the hybrid model fails to be the best method to predict the future government cash needed. The performance of the hybrid model is only superior to the ARIMAX, RBFNN, and ELM methods for both the initial dataset and the selected attributes. Figure 6.34. also provides evidence to the superiority of the deep learning class over the shallow class neural network exhibited based on MASE measurement.

The MASE measurement on the hybrid model supports the finding of Taskaya-Temizel and Casey (2005). The research argued that contrary to its fame, combining several methods into a hybrid model is not always superior to the single method. The rationale for this fact lies in the underlying component of the time series data. As described in Chapter 5., each method has its own strength in capturing the linear or nonlinear patterns of the time series data. The ARIMA based method is well known for its competence in apprehending the linear element of the time series data, while the ANN-based technique is capable of handling the non-linear patterns. In the case of both linear or nonlinear patterns existing in a time series data, the hybrid method is the best procedure to build a forecasting model. The MASE measurement result suggests the best model is the one that is developed using the ANN method. The fact implies that, based on the MASE measurement, the disbursement patterns in

Indonesia tend to be embodied with the nonlinear component. However, an extended discussion on this subject is outside this study's scope.

6.5. Summary

The present study proposed three eminent approaches to forecasting modelling studies namely statistical, machine learning, and the combination of both statistical and machine learning techniques to construct a forecasting model. As described in Chapter 5, each method inherited with merits and handicaps. Statistical based method, such as ARIMA, is well-known for its superiority in capturing the linear component of the data. However, dealing with the nonlinear data is the weakness of ARIMA method. It fails to explain nonlinear elements of the data. Vice versa, with machine learning based method, such as ANN method. While prominence in handling nonlinear patterns in the data, ANN method has a drawback when dealing with the linear elements of the data. Moreover, in the case where both linear and nonlinear patterns exist in the data, the hybrid model is preferable. The study utilised the ARIMAX method as the representation of the statistical approach and ANN method as the representation of the machine learning approach. Some ANN techniques namely FFNN, CFNN, RBFNN, GRNN, ELM, LSTM, and GRU were employed to explore the best ANN method to develop a government cash forecasting model. Lastly, a hybrid model was set up by combining ARIMAX and NARNN methods.

To make sure that the best model was achievable, an attribute selection process was commenced prior to the modelling phase. Through this process, the significant attributes that are influencing the government cash forecasting were identified. The attribute selection process suggested two datasets namely the Initial Dataset and the Selected Attributes. The Initial Dataset allowed all available predictor variables to be included into the modelling phase which was the total daily available fund for intermittent expenditure, F, the day of the week, D, the week of the month, W, the month of the year, M, and policy implementation, $Policy$. On the other hand, the Selected Attributes only involved the significant variables as the predictors in the modelling stage. This study chose the total daily available fund for intermittent expenditure, F, the week of the month, W, the month of the year, M, and policy implementation, $Policy$ as the Selected Attributes.

Moreover, both datasets were fed into the proposed methods such that each forecasting model could be built independently. Hence, in total, there were 18 models developed during the modelling process. Once the models were developed, the performance of each model was measured and compared with the others following six different model performance evaluation criteria in the performance evaluation phase. The six model performance evaluations were MSE, RMSE, MAE, MAPE, SMAPE, and MASE. The best model selected was the one with the smallest score of all proposed performance evaluation criteria.

In summary, the experiment's results show that all performance measurements suggested the best government cash forecasting model was built using the GRU method on the Selected Attributes. Utilising the attribute selection stage on developing a government cash forecasting model, by investigating the significant variables that affect government disbursement patterns, it was proven to be a better procedure compared to employing all attributes. The results show that constructing a forecasting model using the proposed hybrid method failed to be the best technique. This finding is consistent with Taskaya-Temizel and Casey (2005) that combining multiple techniques to construct a forecasting model does not always give the best performance.

Chapter 7 Conclusions

7.1. Introduction

The results and discussions of the experiments conducted in the present study were discussed in detail in Chapter 6. Two research questions with respect to pursuing the objective of this study were developed and discussed in Chapter 4 and Chapter 5. This chapter aims to offer answers to the research questions thus achieving the aims of the study, emphasise the contributions of the present study, review the research limitations and display the opportunities for upcoming research.

The organisation of this chapter is as follows. The answers to the research questions are presented in Section 7.2 followed by the contributions of this research in Section 7.3. The limitations of the present study are discussed in Section 7.4, while the opportunities for future research are argued in Section 7.5. Lastly, Section 7.6. is dedicated as a concluding summary.

7.2. Analysis of Results

The objective of the present study is to develop a government cash forecasting model with accuracy that meets an acceptable level of materiality for the cash manager. This objective can be achieved by answering the primary research aim, which is:

<u>To develop a government cash forecasting model with accuracy that meets an acceptable level of materiality for the cash manager by utilising historical government expenditure data.</u>

Based on previous studies, there are two elements that need to be considered when developing a forecasting model from historical data; identifying relevant data and the methods used to build the model. As a consequence, addressing the aim of developing a government cash forecasting model can be derived from answering the questions related to the data and the methods used to build the model.

As described in Chapter 5, the present study sourced the daily data of expenditures from all spending units in Indonesia for the period of 2009 to 2015 held by the Indonesian cash manager. From all daily data, this study managed to define the intermittent expenditures from the routine expenditures and used the total daily intermittent expenditures as the data to construct a government cash forecasting model. Overall, there are six variables generated from the total daily intermittent expenditures data, including a predictand (total daily intermittent expenditure) and five predictors (total daily available fund for intermittent expenditure, the day of the week, the week of the month, the month of the year, and policy implementation), as detailed in Table 5.1.

To achieve the research aim of this study, the first research question asked:

What are the most appropriate variables to be included in developing a government cash forecasting model with accuracy that meets an acceptable level of materiality for the cash manager by utilising the historical data of government expenditure?

Answering the question, the present study employed the attribute selection process as explained in Chapter 5. Correlations of all six variables are calculated such that the significance of each predictor relative to the predictand is definite. The attribute selection process suggested two sets of variables as a candidate for the best predictor to develop a government cash forecasting model, which are initial dataset and selected attributes. The initial dataset consists of all five initial predictors, while the selected attributes comprise only four predictors (total daily available fund for intermittent expenditure, the week of the month, the month of the year, and policy implementation).

As presented in Chapter 6, the experiments showed that using the same method and measured with the same performance evaluation tool, every model developed with the selected attributes performs better than the one that is built with the initial dataset. Therefore, the best predictors to develop a short-term government cash forecasting model with accuracy that meets an acceptable level of materiality for the cash manager are total daily available fund for intermittent expenditure, the week of the month, the month of the year, and policy implementation.

Furthermore, another aspect to examine regarding the development of a forecasting model from the historical data is the utilised method. Based on previous studies, there are three underlying methods that researchers commonly use to construct a forecasting model: statistical, machine learning, and hybrid models. Each method is embodied with a specific characteristic in relation to its ability in capturing the component of the data. While statistical and machine learning based methods are capable of handling linear and nonlinear elements of the data respectively, the hybrid model is famous for its superiority to detect both linear and nonlinear patterns in the data. In other words, when linear patterns exist in the data, then a statistical-based method is the best technique to develop a government cash forecasting model. In the case where the data follows nonlinear patterns, the machine learning method is the best. If both linear and nonlinear patterns occur in the data, then a hybrid model is the best. All of the patterns' arrangements are possible to occur in the data used in this study.

To address the research aim of this study regarding the method, the second research question asked:

What are the most appropriate techniques to use to develop a government cash forecasting model with accuracy that meets an acceptable level of materiality for the cash manager by utilising historical government expenditure data?

As presented in Chapter 6, the experiments showed that using the same data and measured with the same performance evaluation tool, every model developed with the GRU performs better than the one that is built with other techniques. Since the GRU was developed under the machine learning method, it can be concluded that the total daily intermittent expenditure data, which are used to construct the forecasting model, follow the nonlinear patterns. Therefore, the best technique to develop a short-term government cash forecasting model with accuracy that meets an acceptable level of materiality for the cash manager is the GRU technique.

In conclusion, to address the primary research aim, developing a short-term government cash forecasting model with accuracy that meets an acceptable level of materiality for the cash manager by utilising historical government expenditure data, can be done by utilising the attribute selection process and employing the machine learning method. In the case of the present study, the attribute selection process chooses the total daily available funds for intermittent expenditure, the week of the month, the month of the year, and policy implementation as the best variables to build a government cash forecasting model, while the GRU is selected as the best technique.

7.3. Contribution of the Research

The present study contributes to the body of knowledge in several ways. Firstly, the present study fills the gap in the literature by investigating the procedure to develop a short-term government cash forecasting model with accuracy that meets an acceptable level of materiality for the cash manager by utilising the historical data of government expenditure. The studies that are specifically focussing on developing a government cash forecasting model based on historical data are rare and inconclusive. Almost none of the research agreed on how a government cash forecasting model should be built. The findings from this study confirm that by implementing the attribute selection process and employing the deep learning neural networks technique of ANN method, a short-term government cash forecasting model with accuracy that meets an acceptable level of materiality for the cash manager can be achievable.

Moreover, many studies in the area of prediction modelling emphasise the necessity of understanding the patterns of the historical data to be able to gain a reliable accuracy of the model. The empirical findings in this study provide an insight into the government expenditure patterns. The results provided evidence that the total daily intermittent expenditures patterns, as a proxy to the government expenditure, tend to have nonlinear patterns. A better understanding of the patterns of government expenditure helps the cash manager to establish more effective policies related to the development of government cash forecasting model utilising the historical data of government expenditure.

Furthermore, the study provides a framework for the cash manager to develop a government cash forecasting model by using the historical data of government expenditure. The present study proved that a reliable government cash forecasting model can be accomplished by following three steps: (1) identifying the significant variables that affect the government expenditure through an attribute selection process; (2) exploring the best modelling method through a series of experiments; and (3) evaluating the proposed model to determine the best performance of the government cash forecasting model. Notwithstanding the fact that the present study takes the Indonesian government as an object of study, the procedures presented in this study are applicable to other governments and public sectors.

7.4. Limitations of the Research

The use of Indonesian Government data as the case for this study limits the ability to generalise the reported findings. The variables employed in the present study derived from the Indonesian Government's cash management arrangements which are, by their nature, embodied with specific features of Indonesian Government expenditures. It is logical to think that the characteristics represented in the variables used in this study may not be evidential in other governmental cash management processes, at least to the same extent. The differences in the institutional arrangements and the nature of the economy limit the results of the present study to be able to inform the development of government cash forecasting modelling in other government systems. However, this limitation does not violate the framework contribution of the study. Regardless of the differences, it is possible for other government systems to follow the procedure proposed by this study.

7.5. Opportunities for Future Research

The present study provides evidence of the attribute selection process on prediction performance of the government cash forecasting model. The process aims to identify the predictors that are significantly influencing the future value of government expenditure. There are many other methods that can be used in the attribute selection process. In theory, there are three commonly used attribute selection techniques: filter, wrapper and embedded methods (De Silva & Leong, 2015). Each technique follows a different selection strategy to the observed attributes. This study used the filter-based approach to select the most appropriate variables on developing a government cash forecasting model. Utilising other techniques on the attribute selection process of developing a model for government cash forecasting is a promising area of future research.

Additional areas of potential research value are the choice of the modelling technique. Mu (2006) specifically mentioned analysing the historical data of government expenditure patterns as one of the aspects for consideration to strengthen the capacity for cash forecasting accuracy. The present study provided evidence that the government expenditure patterns followed nonlinear patterns. On the other hand, the machine learning-based method, such as ANN, is famous for its ability to capture the nonlinear patterns in the data. The findings of this study suggested that the GRU of ANN is the best modelling technique to develop a government cash forecasting model. Therefore, research focussing on investigating other ANN techniques as the best method to construct a government cash forecasting model is worth pursuing.

7.6. Conclusion

This study has investigated the best procedure to build a government cash forecasting model. The research aim saw the development of a short-term government cash forecasting model with accuracy that meets an acceptable level of materiality for the cash manager by utilising the historical data of government expenditure. In order to achieve the aim, a series of experiments were conducted with focus on identifying the best predictors and methods to be used.

The conclusion from this research is that a government cash forecasting model with accuracy that meets an acceptable level of materiality for the cash manager can be accomplished by utilising the attribute selection process and employing the machine learning method. In the case of the present study, the attribute selection process chooses the total daily available fund for intermittent expenditure, the week of the month, the month of the year, and policy implementation as the best variables to build a government cash forecasting model, while the GRU is selected as the best technique.

As a contribution, the present study fills the gap in the literature and practice by building on prior studies in identifying procedures and techniques to develop an accurate short-term government cash forecasting model. This study succeeds in that a reliable government cash forecasting model can be attained by utilising the attribute selection process and employing the machine learning method.

The present study also contributes in providing a better understanding of patterns of the government expenditure. The results of this study confirm that government expenditure patterns are non-linear. Insight into patterns of government expenditure help the cash manager to establish more effective policies related to the development of a government cash forecasting model. Moreover, the present study also contributes in providing a framework for the cash manager to develop a government cash forecasting model by using the historical data of government expenditure. However, due to the nature of the data used in the present study, the generality of findings reported in this thesis may be limited. Differences in institutional arrangements and economic structures might hinder the findings from being valid in other government systems. Despite those differences, it is argued that the procedures presented in this study are largely applicable to governments and public sectors.

Furthermore, the present study suggests some areas that might be improved in future research. This study used the filter-based approach for the attribute selection process and employed the ANN-based method to develop a government cash forecasting model. In theory, there are many techniques to select the most appropriate predictors and to capture the nonlinear patterns in the data in order to develop a reliable government cash forecasting model. Therefore, research with a focus on utilising other attribute selection methods and investigating other ANN techniques as the best technique in developing a government cash forecasting model are promising areas of future study.

APPENDICES[4]

Appendix 1 – List of variables of the initial dataset

No	Variable	Description
Dependent Variable		
1	E	The total amount of daily intermittent expenditure (in natural logarithm form).
Independent Variable		
1	F	The total amount of daily remaining budget for intermittent expenditure (in natural logarithm form).
2	$D1$	$D1$ = '1' if the expenditure happens on Monday and '0' otherwise
3	$D2$	$D2$ = '1' if the expenditure happens on Tuesday and '0' otherwise
4	$D3$	$D3$ = '1' if the expenditure happens on Wednesday and '0' otherwise
5	$D4$	$D4$ = '1' if the expenditure happens on Thursday and '0' otherwise
6	$D5$	$D5$ = '1' if the expenditure happens on Friday and '0' otherwise
7	$W1$	$W1$ = '1' if the expenditure happens on the 1st week of the month and '0' otherwise
8	$W2$	$W2$ = '1' if the expenditure happens on the 1st week of the month and '0' otherwise
9	$W3$	$W3$ = '1' if the expenditure happens on the 2nd week of the month and '0' otherwise
10	$W4$	$W4$ = '1' if the expenditure happens on the 3rd week of the month and '0' otherwise
11	$W5$	$W5$ = '1' if the expenditure happens on the 4th week of the month and '0' otherwise
12	$M1$	$M1$ = '1' if the expenditure happens on January and '0' otherwise
13	$M2$	$M2$ = '1' if the expenditure happens on February and '0' otherwise
14	$M3$	$M3$ = '1' if the expenditure happens on March and '0' otherwise
15	$M4$	$M4$ = '1' if the expenditure happens on April and '0' otherwise
16	$M5$	$M5$ = '1' if the expenditure happens on May and '0' otherwise
17	$M6$	$M6$ = '1' if the expenditure happens on June and '0' otherwise
18	$M7$	$M7$ = '1' if the expenditure happens on July and '0' otherwise
19	$M8$	$M8$ = '1' if the expenditure happens on August and '0' otherwise
20	$M9$	$M9$ = '1' if the expenditure happens on September and '0' otherwise
21	$M10$	$M10$ = '1' if the expenditure happens on October and '0' otherwise
22	$M11$	$M11$ = '1' if the expenditure happens on November and '0' otherwise
23	$M12$	$M12$ = '1' if the expenditure happens on December and '0' otherwise
24	$Policy1$	$Policy1$ = '1' if the expenditure happens before 2011 and '0' otherwise
25	$Policy2$	$Policy2$ = '1' if the expenditure happens 2011 and beyond and '0' otherwise

[4] Access to full datasets and detailed experiments is available on request.

Appendix 2 – The comparison of the performance of attribute selection

The datasets used were as follows:

I. total daily available fund for intermittent expenditure, F, the day of the week, D, the week of the month, W, the month of the year, M, and policy implementation, $Policy$
II. total daily available fund for intermittent expenditure, F, the week of the month, W, the month of the year, M, and policy implementation, $Policy$
III. total daily available fund for intermittent expenditure, F, the month of the year, M, and policy implementation, $Policy$
IV. total daily available fund for intermittent expenditure, F, and the month of the year, M
V. total daily available fund for intermittent expenditure, F.

To support the selection of Attribute Selection process, each dataset is used to develop a forecasting model. The selected attributes are the one with the best model's performance. This study employed the ARIMA methods as described in Chapter 5 into the proposed datasets. The comparison of its performance is shown below.

Dataset	ARIMAX's configuration	MSE	RMSE	MAE	MAPE	SMAPE
I	(2,0,3)	1.380469	1.174934	0.725200	2.786200	2.915000
II	**(1,0,2)**	**1.290330**	**1.135927**	**0.605979**	**2.366177**	**2.547979**
III	(1,0,2)	1.392714	1.180133	0.625565	2.443936	2.653045
IV	(1,0,2)	1.443746	1.20156	0.630431	2.465925	2.693263
V	(1,0,2)	1.436409	1.198503	0.619349	2.423552	2.659641

The result shows the best model is the one that is built using dataset II which is employing total daily available fund for intermittent expenditure, F, the week of the month, W, the month of the year, M, and policy implementation, $Policy$ as independent variables. Therefore, it confirms the selected attributes described in Section 6.2.2.

Appendix 3 – List of variables of the selected attributes

No	Variable	Description
Dependent Variable		
1	E	The total amount of daily intermittent expenditure (in natural logarithm form).
Independent Variable		
1	F	The total amount of daily remaining budget for intermittent expenditure (in natural logarithm form).
2	$W1$	$W1$ = '1' if the expenditure happens on the 1st week of the month and '0' otherwise
3	$W2$	$W2$ = '1' if the expenditure happens on the 1st week of the month and '0' otherwise
4	$W3$	$W3$ = '1' if the expenditure happens on the 2nd week of the month and '0' otherwise
5	$W4$	$W4$ = '1' if the expenditure happens on the 3rd week of the month and '0' otherwise
6	$W5$	$W5$ = '1' if the expenditure happens on the 4th week of the month and '0' otherwise
7	$M1$	$M1$ = '1' if the expenditure happens on January and '0' otherwise
8	$M2$	$M2$ = '1' if the expenditure happens on February and '0' otherwise
9	$M3$	$M3$ = '1' if the expenditure happens on March and '0' otherwise
10	$M4$	$M4$ = '1' if the expenditure happens on April and '0' otherwise
11	$M5$	$M5$ = '1' if the expenditure happens on May and '0' otherwise
12	$M6$	$M6$ = '1' if the expenditure happens on June and '0' otherwise
13	$M7$	$M7$ = '1' if the expenditure happens on July and '0' otherwise
14	$M8$	$M8$ = '1' if the expenditure happens on August and '0' otherwise
15	$M9$	$M9$ = '1' if the expenditure happens on September and '0' otherwise
16	$M10$	$M10$ = '1' if the expenditure happens on October and '0' otherwise
17	$M11$	$M11$ = '1' if the expenditure happens on November and '0' otherwise
18	$M12$	$M12$ = '1' if the expenditure happens on December and '0' otherwise
19	$Policy1$	$Policy1$ = '1' if the expenditure happens before 2011 and '0' otherwise
20	$Policy2$	$Policy2$ = '1' if the expenditure happens 2011 and beyond and '0' otherwise

Appendix 4 – Unit Root Test

Null Hypothesis: E has a unit root
Exogenous: Constant
Lag Length: 1 (Automatic - based on SIC, maxlag=22)

		t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic		-13.37553	0.0000
Test critical values:	1% level	-3.435475	
	5% level	-2.863691	
	10% level	-2.567965	

*MacKinnon (1996) one-sided p-values.

Augmented Dickey-Fuller Test Equation
Dependent Variable: D(E)
Method: Least Squares
Date: 11/06/18 Time: 13:03
Sample (adjusted): 1/07/2009 12/31/2013
Included observations: 1225 after adjustments

Variable	Coefficient	Std. Error	t-Statistic	Prob.
E(-1)	-0.281744	0.021064	-13.37553	0.0000
D(E(-1))	-0.214972	0.026824	-8.014108	0.0000
C	7.654454	0.572817	13.36282	0.0000

R-squared	0.229034	Mean dependent var	0.025275
Adjusted R-squared	0.227773	S.D. dependent var	2.354745
S.E. of regression	2.069267	Akaike info criterion	4.294712
Sum squared resid	5232.438	Schwarz criterion	4.307228
Log likelihood	-2627.511	Hannan-Quinn criter.	4.299421
F-statistic	181.5126	Durbin-Watson stat	1.981497
Prob(F-statistic)	0.000000		

Null Hypothesis: F has a unit root
Exogenous: Constant
Lag Length: 0 (Automatic - based on SIC, maxlag=22)

		t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic		-3.851139	0.0025
Test critical values:	1% level	-3.435471	
	5% level	-2.863689	
	10% level	-2.567964	

*MacKinnon (1996) one-sided p-values.

Augmented Dickey-Fuller Test Equation
Dependent Variable: D(F)
Method: Least Squares
Date: 11/06/18 Time: 13:05
Sample (adjusted): 1/06/2009 12/31/2013
Included observations: 1226 after adjustments

Variable	Coefficient	Std. Error	t-Statistic	Prob.
F(-1)	-0.026167	0.006795	-3.851139	0.0001
C	0.867383	0.225485	3.846753	0.0001

R-squared	0.011972	Mean dependent var	-0.000916
Adjusted R-squared	0.011165	S.D. dependent var	0.103051
S.E. of regression	0.102475	Akaike info criterion	-1.716774
Sum squared resid	12.85327	Schwarz criterion	-1.708435
Log likelihood	1054.382	Hannan-Quinn criter.	-1.713636
F-statistic	14.83127	Durbin-Watson stat	1.950159
Prob(F-statistic)	0.000124		

Appendix 5 – Estimation of ARIMA(2,0,3) : Initial Dataset

Dependent Variable: E
Method: ARMA Maximum Likelihood (BFGS)
Sample: 1/05/2009 12/31/2013
Included observations: 1227
Convergence achieved after 152 iterations
Coefficient covariance computed using outer product of gradients

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-320.5367	90.93467	-3.524912	0.0004
F	10.25821	2.715338	3.777877	0.0002
DAY="Monday"	-0.583137	0.200983	-2.901423	0.0038
DAY="Thursday"	-0.112949	0.157218	-0.718419	0.4726
DAY="Tuesday"	-0.463486	0.151877	-3.051726	0.0023
DAY="Wednesday"	-0.189251	0.147041	-1.287057	0.1983
MONTH="August"	14.81002	3.170097	4.671788	0.0000
MONTH="December"	28.82646	3.975524	7.250983	0.0000
MONTH="February"	-7.573080	1.749014	-4.329915	0.0000
MONTH="January"	-11.89639	2.077000	-5.727678	0.0000
MONTH="July"	11.00056	2.787622	3.946217	0.0001
MONTH="June"	7.404312	2.397308	3.088594	0.0021
MONTH="March"	-3.864719	1.303425	-2.965050	0.0031
MONTH="May"	3.592731	1.609189	2.232635	0.0258
MONTH="November"	25.35763	3.793639	6.684250	0.0000
MONTH="October"	21.79146	3.575877	6.094019	0.0000
MONTH="September"	18.18836	3.286046	5.535028	0.0000
WEEK="II"	0.593112	0.400154	1.482210	0.1385
WEEK="III"	1.642870	0.455866	3.603841	0.0003
WEEK="IV"	2.374535	0.509450	4.660981	0.0000
WEEK="V"	3.133450	0.537671	5.827823	0.0000
POLICY="II"	-4.853577	1.214636	-3.995911	0.0001
AR(1)	1.990654	0.003387	587.7778	0.0000
AR(2)	-0.991429	0.003208	-309.0587	0.0000
MA(1)	-1.392879	0.725997	-1.918574	0.0553
MA(2)	0.602953	0.603580	0.998962	0.3180
MA(3)	-0.210073	0.315220	-0.666434	0.5053
SIGMASQ	2.198562	0.112667	19.51373	0.0000

R-squared	0.770452	Mean dependent var	27.04086
Adjusted R-squared	0.765283	S.D. dependent var	3.096065
S.E. of regression	1.499968	Akaike info criterion	3.677759
Sum squared resid	2697.636	Schwarz criterion	3.794421
Log likelihood	-2228.305	Hannan-Quinn criter.	3.721656
F-statistic	149.0488	Durbin-Watson stat	1.982579
Prob(F-statistic)	0.000000		

Inverted AR Roots	1.00+.03i	1.00-.03i	
Inverted MA Roots	1.00	.20+.41i	.20-.41i

Appendix 6 – Estimation of ARIMA(2,0,3) : Selected Attributes

Dependent Variable: EXPENDITURE
Method: ARMA Maximum Likelihood (BFGS)
Sample: 1/05/2009 12/31/2013
Included observations: 1227
Convergence achieved after 12 iterations
Coefficient covariance computed using outer product of gradients

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	97.86733	59.38280	1.648075	0.0996
AVAILABLE_FUND	-2.217731	1.789538	-1.239275	0.2155
MONTH="August"	6.595950	3.355215	1.965880	0.0495
MONTH="December"	13.03360	3.046663	4.277993	0.0000
MONTH="February"	-4.128399	1.636552	-2.522619	0.0118
MONTH="January"	-7.045224	1.960666	-3.593281	0.0003
MONTH="July"	4.974214	3.161625	1.573309	0.1159
MONTH="June"	3.356285	3.140872	1.068584	0.2855
MONTH="March"	-1.947394	1.105078	-1.762223	0.0783
MONTH="May"	1.582070	2.308715	0.685260	0.4933
MONTH="November"	11.23773	3.199485	3.512356	0.0005
MONTH="October"	9.580746	3.274133	2.926194	0.0035
MONTH="September"	7.991455	3.300863	2.421020	0.0156
WEEK="II"	0.109584	0.354578	0.309055	0.7573
WEEK="III"	0.730728	0.354128	2.063459	0.0393
WEEK="IV"	1.006396	0.431758	2.330925	0.0199
WEEK="V"	1.284351	0.472562	2.717845	0.0067
POLICY="II"	-4.549745	1.259690	-3.611798	0.0003
AR(1)	0.982457	0.006608	148.6869	0.0000
MA(1)	-0.359265	0.008828	-40.69638	0.0000
MA(2)	0.241091	0.009449	25.51616	0.0000
SIGMASQ	2.295221	0.026308	87.24491	0.0000

R-squared	0.760360	Mean dependent var	27.04086
Adjusted R-squared	0.756184	S.D. dependent var	3.096065
S.E. of regression	1.528766	Akaike info criterion	3.707263
Sum squared resid	2816.236	Schwarz criterion	3.798927
Log likelihood	-2252.406	Hannan-Quinn criter.	3.741754
F-statistic	182.0660	Durbin-Watson stat	1.982894
Prob(F-statistic)	0.000000		

Inverted AR Roots	.98	
Inverted MA Roots	.18+.46i	.18-.46i

REFERENCES

Abdulrahman, B. M. A., Ahmed, A. Y. A. & Abdellah, A. E. Y. 2018. Forecasting of Sudan Inflation Rates using ARIMA Model. *International Journal of Economics and Financial Issues,* 8**,** 17.

Abrishami, H., Boroujli, M., Amin, M. & Mehrara, M. 2013. Government expenditure and economic growth in Iran. *International Letters of Social and Humanistic Sciences,* 11**,** 76-83.

Acuna, G., Ramirez, C. & Curilem, M. Comparing NARX and NARMAX models using ANN and SVM for cash demand forecasting for ATM. Neural Networks (IJCNN), The 2012 International Joint Conference on, 10-15 June 2012 2012. 1-6.

Adhikari, R. 2015. A neural network based linear ensemble framework for time series forecasting. *Neurocomputing,* 157**,** 231-242.

Adubisi, O., David, I., James, F., Awa, U. & Terna, A. 2018. A predictive Autoregressive Integrated Moving Average (ARIMA) Model for forecasting inflation rates. *Research Journal of Business and Economic Management,* 1**,** 1-8.

Afeef, M., Ihsan, A. & Zada, H. 2018. Forecasting Stock Prices through Univariate ARIMA Modeling.

Akaike, H. 1987. Factor Analysis and AIC. *Psychometrika,* 52**,** 317-332.

Al-Batah, M. S., Alkhasawneh, M. S., Tay, L. T., Ngah, U. K., Hj Lateh, H., Isa, M. & Ashidi, N. 2015. Landslide occurrence prediction using trainable cascade forward network and multilayer perceptron. *Mathematical Problems in Engineering,* 2015.

Al-Hnaity, B. & Abbod, M. 2016. Predicting financial time series data using hybrid model. *Intelligent Systems and Applications.* Springer.

Al-Mahasneh, A. J., Anavatti, S. G. & Garratt, M. A. 2018. Review of Applications of Generalized Regression Neural Networks in Identification and Control of Dynamic Systems. *arXiv preprint arXiv:1805.11236*.

Alam, W., Sinha, K., Kumar, R. R., Ray, M., Rathod, S., Singh, K. & Arya, P. 2018. Hybrid linear time series approach for long term forecasting of crop yield. *Indian Journal of Agricultural Sciences,* 88**,** 1275-1279.

Ali, S., Badar, N. & Fatima, H. 2015. Forecasting production and yield of sugar cane and cotton crops of Pakistan for 2013-2030. *Sarhad Journal of Agriculture,* 31**,** 1-10.

Alkhazaleh, M. M. 2018. Forecasting Banking Volatility in Amman Stock Exchange by Using ARIMA Model. *British Journal of Management,* 29.

Allen, R. & Radev, D. 2010. *Extrabudgetary funds*, International Monetary Fund.

Allen, R. & Tommasi, D. 2001. *Managing Public Expenditure: A Reference Book for Transition Countries*, OECD Publishing.

Alshahrani, M. S. A. & Alsadiq, M. a. J. 2014. *Economic growth and government spending in Saudi Arabia: An empirical investigation*, International Monetary Fund.

Anggraeni, W., Vinarti, R. A. & Kurniawati, Y. D. 2015. Performance comparisons between arima and arimax method in moslem kids clothes demand forecasting: Case study. *Procedia Computer Science,* 72**,** 630-637.

Arabani, S. P. & Komleh, H. E. 2018. The Improvement of Forecasting ATMs Cash Demand of Iran Banking Network Using Convolutional Neural Network. *Arabian Journal for Science and Engineering*, 1-11.

Arghyrou, M. G. & Tsoukalas, J. D. 2011. The Greek Debt Crisis: Likely Causes, Mechanics and Outcomes. *World Economy,* 34**,** 173-191.

Ariyo, A. A., Adewumi, A. O. & Ayo, C. K. Stock price prediction using the ARIMA model. Computer Modelling and Simulation (UKSim), 2014 UKSim-AMSS 16th International Conference, 2014. IEEE, 106-112.

Ayala, H. V. H. & Coelho, L. S. 2016. Cascaded evolutionary algorithm for nonlinear system identification based on correlation functions and radial basis functions neural networks. *Mechanical Systems and Signal Processing,* 68–69**,** 378-393.

Bartlett, Z., Han, L., Nguyen, T. T. & Johnson, P. A Machine Learning Based Approach for the Prediction of Road Traffic Flow on Urbanised Arterial Roads. 2018 IEEE 20th International Conference on High Performance Computing and Communications; IEEE 16th International Conference on Smart City; IEEE 4th International Conference on Data Science and Systems (HPCC/SmartCity/DSS), 2018. IEEE, 1285-1292.

Basu, J. K., Bhattacharyya, D. & Kim, T.-H. 2010. Use of artificial neural network in pattern recognition. *International journal of software engineering and its applications,* 4.

Beale, M. H., Hagan, M. T. & Demuth, H. B. 1992. *Neural Network Toolbox™ User's Guide*, The MathWork, Inc.

Behbahani, H., Amiri, A. M., Imaninasab, R. & Alizamir, M. 2018. Forecasting accident frequency of an urban road network: A comparison of four artificial neural network techniques. *Journal of Forecasting,* 0.

Bengio, Y., Simard, P. & Frasconi, P. 1994. Learning long-term dependencies with gradient descent is difficult. *IEEE transactions on neural networks,* 5**,** 157-166.

Bhandari, R. & Gill, J. 2016. An artificial intelligence ATM forecasting system for hybrid neural networks. *International Journal of Computer Applications,* 133.

Bishop, C. M. 1995. *Neural networks for pattern recognition*, Oxford university press.

Biswas, A., Ghose, M. & Pandit, M. 2014. Comparison of Different Neural Network Architectures for Classification of Feature Transformed Data for Face Recognition. *International Journal of Computer Applications,* 96.

Box, G. E. & Draper, N. R. 1987. *Empirical model-building and response surfaces*, John Wiley & Sons.

Box, G. E. & Jenkins, G. M. 1976. Time series analysis: Forecasting and control. *Holden-Day series in time series analysis.* Holden-Day.

Burnham, K. P. & Anderson, D. R. 2002. *Model selection and multimodel inference: a practical information-theoretic approach*, Springer Science & Business Media.

Butler, D. A. 2006. *Enhancing Automotive Stability Control with Artificial Neural Networks.* PhD, University of Tasmania.

Cadenas, E., Rivera, W., Campos-Amezcua, R. & Heard, C. 2016. Wind speed prediction using a univariate ARIMA model and a multivariate NARX model. *Energies,* 9.

Camelo, H. D. N., Lucio, P. S., Junior, J. B. V. L., De Carvalho, P. C. M. & Dos Santos, D. V. G. 2018. Innovative hybrid models for forecasting time series applied in wind generation based on the combination of time series models with artificial neural networks. *Energy,* 151**,** 347-357.

Campos, J. E. 2001. What is Public Expenditure Management (PEM)? *Asian Development Bank, Governance Brief*.

Catal, C., Fenerci, A., Ozdemir, B. & Gulmez, O. 2015. Improvement of Demand Forecasting Models with Special Days. *Procedia Computer Science,* 59**,** 262-267.

Ceniceros, J. F., Antonanzas-Torres, F., De-Pison, J. M. & Sanz-Garcia, A. 2015. Hybrid Modelling of Multilayer Perceptron Ensembles for Predicting the Response of Bolted Lap Joints. *Logic Journal of the IGPL***,** 451-462.

Chaâbane, N. 2014. A hybrid ARFIMA and neural network model for electricity price prediction. *International Journal of Electrical Power & Energy Systems,* 55**,** 187-194.

Chakravorty, J., Shah, S. & Nagraja, H. 2018. ANN and ANFIS for Short Term Load Forecasting. *Engineering, Technology & Applied Science Research,* 8**,** 2818-2820.

Chen, L., Yang, L., Du, J., Sun, C., Du, S. & Xi, H. 2017. An Extreme Learning Machine Architecture Based on Volterra Filtering and PCA. *IEICE Transactions on Information and Systems,* 100**,** 2690-2701.

Chen, X., Dong, Z. Y., Meng, K., Xu, Y., Wong, K. P. & Ngan, H. W. 2012. Electricity Price Forecasting With Extreme Learning Machine and Bootstrapping. *IEEE Transactions on Power Systems,* 27**,** 2055-2062.

Chin, C. S., Ji, X., Woo, W. L., Kwee, T. J. & Yang, W. 2017. Modified multiple generalized regression neural network models using fuzzy C-means with principal component analysis for noise prediction of offshore platform. *Neural Computing and Applications*.

Chinnathambi, R. A., Mukherjee, A., Campion, M., Salehfar, H., Hansen, T., Lin, J. & Ranganathan, P. 2018. A Multi-Stage Price Forecasting Model for Day-Ahead Electricity Markets. *Forecastinsr,* 1**,** 3.

Chipaumire, G., Ngirande, H. & Ruswa, Y. 2014. The Impact of Government Spending on Economic Growth: Case South Africa. *Mediterranean Journal of Social Sciences,* 5**,** 109.

Cho, K., Van Merriënboer, B., Gulcehre, C., Bahdanau, D., Bougares, F., Schwenk, H. & Bengio, Y. 2014. Learning phrase representations using RNN encoder-decoder for statistical machine translation. *arXiv preprint arXiv:1406.1078*.

Chollet, F. 2015. Keras. GitHub.

Chude, N. P. & Chude, D. I. 2013. Impact of government expenditure on economic growth in Nigeria. *International journal of business and management review,* 1**,** 64-71.

Chung, J., Gulcehre, C., Cho, K. & Bengio, Y. 2014. Empirical evaluation of gated recurrent neural networks on sequence modeling. *arXiv preprint arXiv:1412.3555*.

Cogoljević, D., Alizamir, M., Piljan, I., Piljan, T., Prljić, K. & Zimonjić, S. 2018. A machine learning approach for predicting the relationship between energy resources and economic development. *Physica A: Statistical Mechanics and its Applications,* 495**,** 211-214.

Cohen, J. 1988. *Statistical power analysis for the behavioral sciences,* New Jersey, Lawrence Erlbaum Associates.

Comptroller and Auditor General 2014. Forecasting in government to achieve value for money. *In:* MORSE, A. (ed.). London: The National Audit Office.

Cools, M., Moons, E. & Wets, G. Investigating variablility in daily traffic counts using ARIMAX and SARIMA(X) models: assessing impact of holidays on two divergent site locations. TRB 88th Annual Meeting Compendium of Papers DVD, 2009.

Cotterell, P. & Wickens, T. 2007. *SELECTION OF DATA SOURCES, PREPARATIONS, AND THE COMPILATION OF GFSM 2001 STATISTICS*, International Monetary Fund.

Dandekar, P. V. & Ranade, K. M. 2015. ATM Cash Flow Management. *International Journal of Innovation, Management and Technology,* 6**,** 343.

De Myttenaere, A., Golden, B., Le Grand, B. & Rossi, F. 2016. Mean absolute percentage error for regression models. *Neurocomputing,* 192**,** 38-48.

De Oliveira, E. M. & Oliveira, F. L. C. 2018. Forecasting mid-long term electric energy consumption through bagging ARIMA and exponential smoothing methods. *Energy,* 144**,** 776-788.

De Oliveira, J. F. L. & Ludermir, T. B. 2016. A hybrid evolutionary decomposition system for time series forecasting. *Neurocomputing,* 180**,** 27-34.

De Silva, A. M. & Leong, P. H. 2015. *Grammar-based feature generation for time-series prediction*, Springer.

Dg of Treasury 2007. Regulation of The Directure General of Treasury Number PER-35/PB/2007 on the guidelines on government accounts management. *In:* MINISTRY OF FINANCE (ed.).

Dg of Treasury 2013. Internal Memo Number SE-41/PB/2013 on the list of name and code of the treasury office. Jakarta: Directorate General of Treasury.

Diamond, J. & Khemani, P. 2005. *Introducing financial management information systems in developing countries*, International Monetary Fund.

Dickey & Fuller 1979.

Ding, S., Xu, X. & Nie, R. 2014. Extreme learning machine and its applications. *Neural Computing and Applications,* 25**,** 549-556.

Ding, X., Zhang, Y., Liu, T. & Duan, J. Deep learning for event-driven stock prediction. Twenty-Fourth International Joint Conference on Artificial Intelligence, 2015.

Donahue, J., Anne Hendricks, L., Guadarrama, S., Rohrbach, M., Venugopalan, S., Saenko, K. & Darrell, T. Long-term recurrent convolutional networks for visual recognition and description. Proceedings of the IEEE conference on computer vision and pattern recognition, 2015. 2625-2634.

Dong, Y., Li, S. & Gong, X. Time Series Analysis: An application of ARIMA model in stock price forecasting. 2017 International Conference on Innovations in Economic Management and Social Science (IEMSS 2017), 2017. Atlantis Press.

Dudek, G. 2016. Multilayer perceptron for GEFCom2014 probabilistic electricity price forecasting. *International Journal of Forecasting,* 32**,** 1057-1060.

Ebtehaj, I. & Bonakdari, H. 2016. Bed load sediment transport estimation in a clean pipe using multilayer perceptron with different training algorithms. *KSCE Journal of Civil Engineering,* 20**,** 581-589.

Eid, H. F., Hassanien, A. E., Kim, T.-H. & Banerjee, S. 2013. Linear correlation-based feature selection for network intrusion detection model. *Advances in Security of Information and Communication Networks.* Springer.

Feng, Y., Cui, N., Gong, D., Zhang, Q. & Zhao, L. 2017. Evaluation of random forests and generalized regression neural networks for daily reference evapotranspiration modelling. *Agricultural Water Management,* 193**,** 163-173.

Frank, E., Hall, M. A. & Witten, I. H. 2016. *Data Mining: Practical Machine Learning Tools and Techniques*, Morgan Kaufmann.

Fu, R. 2016. Using LSTM and GRU neural network methods for traffic flow prediction. IEEE.

Gamboa, J. C. B. 2017. Deep learning for time-series analysis. *arXiv preprint arXiv:1701.01887*.

Gavin, H. 2017. The Levenberg-Marquardt method for nonlinear least squares curve-fitting problems. Department of Civil and Environmental Engineering, Duke University.

Ghasemiyeh, R., Moghdani, R. & Sana, S. S. 2017. A hybrid artificial neural network with metaheuristic algorithms for predicting stock price. *Cybernetics and Systems,* 48**,** 365-392.

Ghiasi, M. 2014. The Role of Public Expenditure Management in the Budget Systems. *International SAMANM Journal of Marketing and Management,* 2.

Government of Indonesia 2004. Government Regulation Number 21 of 2004 on The Management of State/Regional Money. *In:* GOVERNMENT, T. I. (ed.). Jakarta.

Government of Indonesia 2007. Government Regulation Number 39 of 2007 on The Management of State/Regional Money. *In:* GOVERNMENT, T. I. (ed.). Jakarta.

Granger, C. W. & Terasvirta, T. 1993. Modelling non-linear economic relationships. *OUP Catalogue*.

Graves, A. 2013. Generating sequences with recurrent neural networks. *arXiv preprint arXiv:1308.0850*.

Graves, A. & Jaitly, N. Towards end-to-end speech recognition with recurrent neural networks. International Conference on Machine Learning, 2014. 1764-1772.

Gujarati, D. N. & Porter, D. C. 2009. *Basic Econometrics,* Boston, McGraw-Hill/Irwin.

Guyon, I. 2008. Practical feature selection: from correlation to causality. *Mining massive data sets for security: advances in data mining, search, social networks and text mining, and their applications to security***,** 27-43.

Guyon, I. & Elisseeff, A. 2003. An introduction to variable and feature selection. *Journal of machine learning research,* 3**,** 1157-1182.

Hall, M., Frank, E., Holmes, G., Pfahringer, B., Reutemann, P. & Witten, I. H. 2009. The WEKA data mining software: an update. *SIGKDD explorations,* 11**,** 10-18.

Hall, M. A. 1999. Correlation-based feature selection for machine learning.

Hannan, E. J. & Quinn, B. G. 1979. The determination of the order of an autoregression. *Journal of the Royal Statistical Society. Series B (Methodological)***,** 190-195.

Harrell Jr, F. E. 2015. *Regression Modeling Strategies with Applications to Linear Models, Logistic and Ordinal Regression and Survival Analysis,* Heidelberg, Springer-Verlag.

Haykin, S. 1999. *Neural Networks: A Comprehensive Foundation,* USA, Prentice Hall International, Inc.

Haykin, S. 2009. *Neural Networks and Learning Machines,* USA, Prentice Hall International, Inc.

Heinze, G., Wallisch, C. & Dunkler, D. 2018. Variable selection–A review and recommendations for the practicing statistician. *Biometrical Journal,* 60**,** 431-449.

Hernandez, L. & Montiel, P. J. 2003. Post-crisis exchange rate policy in five Asian countries: Filling in the “hollow middle”? *Journal of the Japanese and International Economies,* 17**,** 336-369.

Hiransha, M., Gopalakrishnan, E., Menon, V. K. & Soman, K. 2018. NSE stock market prediction using deep-learning models. *Procedia computer science,* 132**,** 1351-1362.

Hochreiter, S. 1991. Untersuchungen zu dynamischen neuronalen Netzen. *Diploma, Technische Universität München,* 91.

Hochreiter, S. & Schmidhuber, J. 1997. Long short-term memory. *Neural computation,* 9**,** 1735-1780.

Hosseini, S. M. H. & Shabanian, M. A New Hybrid Intelligent Approach for Traffic Flow Forecasting based on Fuzzy Controllers. IECON 2018-44th Annual Conference of the IEEE Industrial Electronics Society, 2018. IEEE, 2177-2182.

Huang, G.-B., Zhu, Q.-Y. & Siew, C.-K. Extreme learning machine: a new learning scheme of feedforward neural networks. 2004 IEEE International Joint Conference on Neural Networks, 2004 Budapest, Hungary. IEEE, 985.

Huang, G.-B., Zhu, Q.-Y. & Siew, C.-K. 2006. Extreme learning machine: theory and applications. *Neurocomputing,* 70**,** 489-501.

Huang, G., Huang, G.-B., Song, S. & You, K. 2015. Trends in extreme learning machines: A review. *Neural Networks,* 61**,** 32-48.

Huang, H. Analysis and Forecast of Resource-based City GDP Based on ARIMA Model. 2019 4th International Conference on Financial Innovation and Economic Development (ICFIED 2019), 2019. Atlantis Press.

Hyndman, R. J. & Athanasopoulos, G. 2018. *Forecasting: principles and practice*, OTexts.

Hyndman, R. J. & Koehler, A. B. 2006. Another look at measures of forecast accuracy. *International journal of forecasting,* 22**,** 679-688.

Ihs Global 2016. *Eviews 9 User's Guide II,* Irvine, IHS Global Inc.

Ilić, I., Jovanović, S. & Janković–Milić, V. 2016. Forecasting corn production in Serbia using Arima model. *Economics of Agriculture,* 63**,** 1141-1156.

Imf 2014. *Government Finance Statistics Manual 2014,* Washington, D.C.

Ince, H. & TrafaliS, T. B. 2017. A hybrid forecasting model for stock market prediction. *Economic Computation & Economic Cybernetics Studies & Research,* 51.

Inthachot, M., Boonjing, V. & Intakosum, S. 2016. Artificial neural network and genetic algorithm hybrid intelligence for predicting thai stock price index trend. *Computational intelligence and neuroscience,* 2016.

Iqbal, M. & Naveed, A. 2016. Forecasting inflation: Autoregressive integrated moving average model. *European Scientific Journal,* 12.

Iskandar, I., Wiillet, R. & Xu, S. 2018. The development of a government cash forecasting model. *Journal of Public Budgeting, Accounting & Financial Management,* 30**,** 368-383.

Jain, A. K., Mao, J. & Mohiuddin, K. 1996. Artificial neural networks: A tutorial. *Computer***,** 31-44.

Jang, H. & Lee, J. 2018. An empirical study on modeling and prediction of bitcoin prices with bayesian neural networks based on blockchain information. *IEEE Access,* 6**,** 5427-5437.

Ji, J. & Hou, J. Forecast on Bus Trip Demand Based on ARIMA Models and Gated Recurrent Unit Neural Networks. 2017 International Conference on Computer Systems, Electronics and Control (ICCSEC), 2017. IEEE, 105-108.

Jozefowicz, R., Zaremba, W. & Sutskever, I. An empirical exploration of recurrent network architectures. International Conference on Machine Learning, 2015. 2342-2350.

Keynes, J. M. 1936. *The general theory of employment, interest, and money,* New York, Harcourt, Brace and Co.

Khan, K., Khan, G., Shaikh, S. A., Lodhi, A. S. & Jilani, G. 2015. ARIMA Modelling for Forecasting of Rice Production: A Case Study of Pakistan. *Lasbela University Journal of Sciences and Technology,* 4**,** 117-120.

Khashei, M., Bijari, M. & Raissi Ardali, G. A. 2009. Improvement of Auto-Regressive Integrated Moving Average models using Fuzzy logic and Artificial Neural Networks (ANNs). *Neurocomputing,* 72**,** 956-967.

Khashei, M. & Hajirahimi, Z. 2018. A comparative study of series arima/mlp hybrid models for stock price forecasting. *Communications in Statistics-Simulation and Computation***,** 1-16.

Khemani, P. & Diamond, J. 2005. *Introducing financial Management Information Systems in Developing Countries (EPub)*, International Monetary Fund.

Kim, S. & Kim, H. 2016. A new metric of absolute percentage error for intermittent demand forecasts. *International Journal of Forecasting,* 32**,** 669-679.

Kouretas, G. P. & Vlamis, P. 2010. The Greek crisis: causes and implications. *Panoeconomicus,* 57**,** 391-404.

Kuckuck, J. 2012. Testing Wagner's law at different stages of economic development: A historical analysis of five Western European countries. *Working Paper.* Institute of Empirical Economic Research, University of Osnabrück.

Kumar, S. V. & Vanajakshi, L. 2015. Short-term traffic flow prediction using seasonal ARIMA model with limited input data. *European Transport Research Review,* 7**,** 21.

Lamartina, S. & Zaghini, A. 2011. Increasing Public Expenditure: Wagner's Law in OECD Countries. *German Economic Review,* 12**,** 149-164.

Lane, T. D., Schulze-Gattas, M., Tsikata, T., Phillips, S., Ghosh, A. R. & Hamann, A. J. 1999. *IMF-supported programs in Indonesia, Korea and Thailand*, International Monetary Fund.

Levenberg, K. 1944. A method for the solution of certain non-linear problems in least squares. *Quarterly of applied mathematics,* 2**,** 164-168.

Li, K.-L., Zhai, C.-J. & Xu, J.-M. Short-term traffic flow prediction using a methodology based on ARIMA and RBF-ANN. 2017 Chinese Automation Congress (CAC), 2017. IEEE, 2804-2807.

Li, K., Xie, X., Xue, W., Dai, X., Chen, X. & Yang, X. 2018. A hybrid teaching-learning artificial neural network for building electrical energy consumption prediction. *Energy and Buildings,* 174**,** 323-334.

Lienert, I. 2008. Cash Management. *IMF Fiscal Affairs Department, Public Financial Management Technical Guidance Note*.

Lienert, I. 2009. Modernizing Cash Management. *IMF Technical Notes and Manuals (Washington: International Monetary Fund)*.

Lolli, F., Gamberini, R., Regattieri, A., Balugani, E., Gatos, T. & Gucci, S. 2017. Single-hidden layer neural networks for forecasting intermittent demand. *International Journal of Production Economics,* 183**,** 116-128.

Lutu, P. E. N. 2018. *Dataset selection for aggregate model implementation in predictive data mining.*

Lv, Y., Duan, Y., Kang, W., Li, Z. & Wang, F.-Y. 2015. Traffic flow prediction with big data: A deep learning approach. *IEEE Trans. Intelligent Transportation Systems,* 16**,** 865-873.

Magazzino, C. 2011. The nexus between public expenditure and inflation in the Mediterranean countries. *Theoretical and Practical Research in Economic Fields (TPREF)***,** 94-107.

Magazzino, C. 2012. Wagner versus Keynes: Public spending and national income in Italy. *Journal of Policy Modeling,* 34**,** 890-905.

Marquardt, D. W. 1963. An algorithm for least-squares estimation of nonlinear parameters. *Journal of the society for Industrial and Applied Mathematics,* 11**,** 431-441.

Meier, J.-H., Schneider, S., Schmidt, I., Schüller, P., Schönfeldt, T. & Wanke, B. ANN-Based Electricity Price Forecasting Under Special Consideration of Time Series Properties. International Conference on Information and Communication Technologies in Education, Research, and Industrial Applications, 2018. Springer, 262-275.

Menyah, K. & Wolde-Rufael, Y. 2013. Government Expenditure and Economic Growth: The Ethiopian Experience, 1950-2007. *Journal of Developing Areas,* 47**,** 263-280.

Micchelli, C. A. 1984. Interpolation of scattered data: distance matrices and conditionally positive definite functions. *Approximation theory and spline functions.* Springer.

Miller, J. 2018. ARIMA Time Series Models for Full Truckload Transportation Prices. *Forecasting,* 1**,** 121-134.

Minister of Finance 2004. Presidential Decree number 36 Year 2004 concerning Revision of Presidential Decree number 109 Year 2001 concerning Organisation and Duty of Echelon I of Ministry of Republic of Indonesia. *In:* FINANCE, M. O. (ed.). Jakarta.

Minister of Finance 2006a. Finance Ministerial Regulation number 68/PMK.06/2006 concerning Piloting of Zero Balanced Spending Account in Treasury Office's Bank Partner on Implemantation of TSA. *In:* FINANCE, M. O. (ed.). Jakarta.

Minister of Finance 2006b. Presidential Regulation number 66 Year 2006 concerning Fourth Revision of Presidential Regulation number 10 Year 2005 concerning Organisation and Duty of Echelon I of Ministry of Republic of Indonesia. *In:* FINANCE, M. O. (ed.). Jakarta.

Minister of Finance 2007a. Finance Ministerial Regulation number 58/PMK.05/2007 concerning The arrangement of government accounts within ministries/agencies. *In:* FINANCE, M. O. (ed.). Jakarta.

Minister of Finance 2007b. Finance Ministerial Regulation number 98/PMK.05/2007 concerning Implementation of Zero Balanced Spending Account in Treasury Office on Implemantation of TSA. *In:* FINANCE, M. O. (ed.). Jakarta.

Minister of Finance 2009a. Finance Ministerial Regulation Number 61/KMK.05/2009 Concerning: Implementation of Treasury Notional Pooling in the Spending Treasurer Account. *In:* MINISTRY OF FINANCE (ed.).

Minister of Finance 2009b. Finance Ministerial Regulation number 116/PMK.05/2009 concerning Piloting of Zero Balanced Revenue Account on Implemantation of TSA *In:* FINANCE, M. O. (ed.). Jakarta.

Minister of Finance 2009c. Finance Ministerial Regulation number 192/PMK.05/2009 concerning Cash Planning. *In:* FINANCE, M. O. (ed.). Jakarta.

Minister of Finance 2009d. Joint Decree of Minister of Finance and Governor of Bank Indonesia Number 17/KMK.05/2009 and Number 11/3/KEP.GBI/2009 Concerning: Coordination of Government Cash Management The Minister of Finance and The Governor of Bank Indonesia. *In:* MINISTRY OF FINANCE & BANK INDONESIA (eds.).

Minister of Finance 2009e. Regulation of The Finance Minister Number 126/KMK.05/2009 on The Application of Treasury Notional Pooling in the Receiving Treasurer's Account. *In:* MINISTRY OF FINANCE (ed.).

Minister of Finance 2010a. Finance Ministerial Regulation number 3/PMK.05/2010 concerning Management of Government Cash Excess/Shortage. *In:* FINANCE, M. O. (ed.). Jakarta.

Minister of Finance 2010b. Finance Ministerial Regulation number 32/PMK.05/2010 concerning Implementation of Zero Balanced Revenue Account on Implemantation of TSA *In:* FINANCE, M. O. (ed.). Jakarta.

Minister of Finance 2010c. Speech at The Handover Ceremony of The Budget Execution Document. *In:* MINISTRY OF FINANCE (ed.). Jakarta.

Minister of Finance. 2011. *History of State Treasury and Budget System* [Online]. Jakarta: Ministry of Finance. Available: http://www.span.depkeu.go.id/content/sejarah/ [Accessed 23 March 2016 2016].

Minister of Finance 2013. Finance Ministerial Regulation number 154/PMK.05/2013 concerning Piloting of State Treasury and Budget System. *In:* FINANCE, M. O. (ed.). Jakarta.

Minister of Finance 2014a. Finance Ministerial Regulation number 154/PMK.05/2014 concerning Implementation of State Treasury and Budget System. *In:* FINANCE, M. O. (ed.). Jakarta.

Minister of Finance 2014b. Finance Ministerial Regulation number 277/PMK.05/2014 concerning Disbursement Plan, Receipts Plan, and Cash Planning. *In:* FINANCE, M. O. (ed.). Jakarta.

Minister of Finance 2015a. Finance Ministerial Regulation number 127/PMK.02/2015 concerning Budget Classification. *In:* INDONESIA, M. O. F. O. (ed.). Jakarta.

Minister of Finance 2015b. Regulation of The Finance Minister Number 223/PMK.05/2015 on The Piloting of Spending Units' Financial Application System. *In:* MINISTRY OF FINANCE (ed.).

Minister of Finance 2016. Finance Ministerial Regulation number 115/PMK.05/2016 concerning revision of Finance Ministerial Regulation number 3/PMK.05/2010 concerning Management of Government Cash Excess/Shortage. *In:* FINANCE, M. O. (ed.). Jakarta.

Mishra, P. S. & Dehuri, S. 2014. Potential Indicators Based Neural Networks for Cash Forecasting of an ATM. *International Journal of Information Systems and Social Change (IJISSC),* 5**,** 41-57.

Modaresi, F., Araghinejad, S. & Ebrahimi, K. 2018. A Comparative Assessment of Artificial Neural Network, Generalized Regression Neural Network, Least-Square Support Vector Regression, and K-Nearest Neighbor Regression for Monthly Streamflow Forecasting in Linear and Nonlinear Conditions. *Water Resources Management,* 32**,** 243-258.

Moghaddam, A. H., Moghaddam, M. H. & Esfandyari, M. 2016. Stock market index prediction using artificial neural network. *Journal of Economics, Finance and Administrative Science,* 21**,** 89-93.

Mondal, P., Shit, L. & Goswami, S. 2014. Study of effectiveness of time series modeling (ARIMA) in forecasting stock prices. *International Journal of Computer Science, Engineering and Applications,* 4**,** 13.

Montgomery, D. C., Jennings, C. L. & Kulahci, M. 2015. *Introduction to time series analysis and forecasting*, John Wiley & Sons.

Moretti, F., Pizzuti, S., Panzieri, S. & Annunziato, M. 2015. Urban traffic flow forecasting through statistical and neural network bagging ensemble hybrid modeling. *Neurocomputing,* 167**,** 3-7.

Mu, Y. 2006. Government Cash Management: Good Practice & Capacity-Building Framework. *World Bank Financial Discussion Series*.

Nemeshaev, S. & Tsyganov, A. 2016. Model of the Forecasting Cash Withdrawals in the ATM Network. *Procedia Computer Science,* 88**,** 463-468.

Neshat, N., Hadian, H. & Behzad, M. 2018. Nonlinear ARIMAX model for long–term sectoral demand forecasting. *Management Science Letters,* 8**,** 581-592.

Nyoni, T. 2018. Box-Jenkins ARIMA approach to predicting net FDI inflows in Zimbabwe.

Ogujiuba, K. K. & Ehigiamusoe, K. 2014. Capital budget implementation in Nigeria: Evidence from the 2012 capital budget. *Contemporary Economics,* 8**,** 299-314.

Oluwatobi, S. O. & Ogunrinola, O. I. 2011. Government expenditure on human capital development: Implications for economic growth in Nigeria. *Journal of sustainable development,* 4**,** 72.

Panchal, G., Ganatra, A., Kosta, Y. & Panchal, D. 2011. Behaviour analysis of multilayer perceptronswith multiple hidden neurons and hidden layers. *International Journal of Computer Theory and Engineering,* 3**,** 332.

Panda, B. N., Bahubalendruni, M. V. a. R. & Biswal, B. B. 2015. A general regression neural network approach for the evaluation of compressive strength of FDM prototypes. *Neural Computing and Applications,* 26**,** 1129-1136.

Pascanu, R., Gulcehre, C., Cho, K. & Bengio, Y. 2013. How to construct deep recurrent neural networks. *arXiv preprint arXiv:1312.6026*.

Pattanayak, S. & Fainboim, I. 2010. *Treasury Single Account: Concept, Design, and Implementation Issues*, International Monetary Fund.

Pessoa, M. & Williams, M. 2012. *Government Cash Management: Relationship between the Treasury and the Central Bank*, International Monetary Fund.

Pontius, R. G., Thontteh, O. & Chen, H. 2008. Components of information for multiple resolution comparison between maps that share a real variable. *Environmental and Ecological Statistics,* 15**,** 111-142.

Potter, B. H. & Diamond, J. 1999. *Guidelines for public expenditure management*, International Monetary Fund.

Powers, D. & Xie, Y. 2008. *Statistical methods for categorical data analysis*, Emerald Group Publishing.

Prema, V. & Rao, K. U. 2015. Development of statistical time series models for solar power prediction. *Renewable Energy,* 83**,** 100-109.

Rahim, N. F., Othman, M. & Sokkalingam, R. A Comparative Review on Various Method of Forecasting Crude Palm Oil Prices. Journal of Physics: Conference Series, 2018. IOP Publishing, 012043.

Rathod, S., Singh, K., Patil, S., Naik, R. H., Ray, M. & Meena, V. S. 2018. Modeling and forecasting of oilseed production of India through artificial intelligence techniques. *INDIAN JOURNAL OF AGRICULTURAL SCIENCES,* 88**,** 22-27.

Reiner, P. D. 2015. *Algorithms for Optimal Construction and Training of Radial Basis Function Neural Networks.* Auburn University.

Ristolainen, K. 2018. Predicting banking crises with artificial neural networks: The role of nonlinearity and heterogeneity. *The Scandinavian Journal of Economics,* 120**,** 31-62.

Rodin-Brown, E. 2008. Integrated Financial Management Information Systems: A Practical Guide.

Royston, P. & Sauerbrei, W. 2008. *Multivariable model-building: a pragmatic approach to regression anaylsis based on fractional polynomials for modelling continuous variables*, John Wiley & Sons.

Ruiz, L. G. B., Cuéllar, M. P., Calvo-Flores, M. D. & Jiménez, M. D. C. P. 2016. An Application of Non-Linear Autoregressive Neural Networks to Predict Energy Consumption in Public Buildings. *Energies,* 9**,** 684.

Sajjadi, S., Shamshirband, S., Alizamir, M., Yee, P. L., Mansor, Z., Manaf, A. A., Altameem, T. A. & Mostafaeipour, A. 2016. Extreme learning machine for prediction of heat load in district heating systems. *Energy and Buildings,* 122**,** 222-227.

Salman, N., Lawi, A. & Syarif, S. Artificial Neural Network Backpropagation with Particle Swarm Optimization for Crude Palm Oil Price Prediction. Journal of Physics: Conference Series, 2018. IOP Publishing, 012088.

Samudram, M., Nair, M. & Vaithilingam, S. 2009. Keynes and Wagner on Government Expenditures and Economic Development: The Case of a Developing Economy. *Empirical Economics,* 36**,** 697-712.

Schiavo-Campo, S. 2017. *Government Budgeting and Expenditure Management: Principles and International Practice*, Taylor & Francis.

Schick, A. 1998. A contemporary approach to public expenditure management. *World Bank Institute*.

Schmidhuber, J. 2015. Deep learning in neural networks: An overview. *Neural networks,* 61**,** 85-117.

Schwarz, G. 1978. Estimating the dimension of a model. *The annals of statistics,* 6**,** 461-464.

Sharma, B., Kumar, S., Tiwari, P., Yadav, P. & Nezhurina, M. I. 2018. ANN based short-term traffic flow forecasting in undivided two lane highway. *Journal of Big Data,* 5**,** 48.

Siswanto, A. D. & Rahayu, S. L. 2010. *The Contributing Factors of the Low Absorption Ministries/Agencies' Spending in 2010* [Online]. Available:

http://www.kemenkeu.go.id/Artikel/faktor-faktor-penyebab-rendahnya-penyerapan-belanja-kementerianlembaga-ta-2010 [Accessed 02-02-2017].

Specht, D. F. 1991. A general regression neural network. *IEEE Transactions on Neural Networks,* 2**,** 568-576.

Staub, S., Karaman, E., Kaya, S., Karapınar, H. & Güven, E. 2015. Artificial Neural Network and Agility. *Procedia - Social and Behavioral Sciences,* 195**,** 1477-1485.

Steyerberg, E. W. 2008. *Clinical prediction models: a practical approach to development, validation, and updating*, Springer Science & Business Media.

Storkey, I. 2003. Government Cash and Treasury Management Reform. *Asian Development Bank, Governance Brief*.

Sulistyowati, R., Kuswanto, H. & Astuti, E. T. Hybrid forecasting model to predict air passenger and cargo in Indonesia. 2018 International Conference on Information and Communications Technology (ICOIACT), 2018. IEEE, 442-447.

Sumando, E., Hamiyani, F. D. & Purba, I. D. 2018. Pengembangan Metode Cash Forecasting Pemerintah: Studi Kasus Saldo Kas Pemerintah 2009–2011. *Kajian Ekonomi dan Keuangan,* 2**,** 70-93.

Tang, T. C. 2010. Wagner's Law Versus Keynesian Hypothesis in Malaysia: An Impressionistic View. *International Journal of Business and Society,* 11**,** 87.

Taskaya-Temizel, T. & Casey, M. C. 2005. A comparative study of autoregressive neural network hybrids. *Neural Networks,* 18**,** 781-789.

Tealab, A., Hefny, H. & Badr, A. 2018. Forecasting of nonlinear time series using artificial neural network. *Future Computing and Informatics Journal,* 3**,** 143-151.

Thakur, G. S. M., Bhattacharyya, R. & Mondal, S. S. 2016. Artificial neural network based model for forecasting of inflation in India. *Fuzzy Information and Engineering,* 8**,** 87-100.

The Pennsylvania State University. 2018. *STAT 510 – Applied Time Series Analysis* [Online]. Available: https://newonlinecourses.science.psu.edu/stat510/ [Accessed 26 December 2018].

Ülke, V., Sahin, A. & Subasi, A. 2018. A comparison of time series and machine learning models for inflation forecasting: empirical evidence from the USA. *Neural Computing and Applications,* 30**,** 1519-1527.

Usmani, M., Ebrahim, M., Adil, S. H. & Raza, K. Predicting Market Performance with Hybrid Model. 2018 3rd International Conference on Emerging Trends in Engineering, Sciences and Technology (ICEEST), 2018. IEEE, 1-4.

Venkatesh, K., Ravi, V., Prinzie, A. & Van Den Poel, D. 2014. Cash demand forecasting in ATMs by clustering and neural networks. *European Journal of Operational Research,* 232**,** 383-392.

Wagner, A. 1883. *Finanzwissenschaft,* Leipzig, CF Winter.

Wallström, P. 2009. *Evaluation of forecasting techniques and forecast errors: with focus on intermittent demand.* Luleå tekniska universitet.

Wang, B., Liu, P., Chao, Z., Junmei, W., Chen, W., Cao, N., O'hare, G. M. & Wen, F. 2018. Research on Hybrid Model of Garlic Short-term Price Forecasting based on Big Data. *CMC-COMPUTERS MATERIALS & CONTINUA,* 57**,** 283-296.

Wang, L., Zou, H., Su, J., Li, L. & Chaudhry, S. 2013. An ARIMA-ANN Hybrid Model for Time Series Forecasting. *Systems Research and Behavioral Science,* 30**,** 244-259.

Wang, W.-C., Chau, K.-W., Qiu, L. & Chen, Y.-B. 2015. Improving forecasting accuracy of medium and long-term runoff using artificial neural network based on EEMD decomposition. *Environmental Research,* 139**,** 46-54.

Warsito, B., Santoso, R. & Yasin, H. Cascade Forward Neural Network for Time Series Prediction. Journal of Physics: Conference Series, 2018. IOP Publishing, 012097.

Weerathunga, H. & Silva, A. DRNN-ARIMA Approach to Short-term Trend Forecasting in Forex Market. 2018 18th International Conference on Advances in ICT for Emerging Regions (ICTer), 2018. IEEE, 287-293.

Wei, W. W. S. 2006. *Time Series Analysis: Univariate and Multivariate Methods*, Pearson Addison Wesley.

Widodo, R., Sihombing, W. P., Budhisusetyo, A. & Harjowiryono, M. 2014. *Cash Management Reform in Indonesia : Making The State Money Work Harder,* Washington, DC, World Bank Group.

Williams, M. 2004. Government Cash Management: Good and Bad Practice. *World Bank Technical Note*.

Williams, M. 2009. Government Cash Management: International Practice. Oxford Policy Management Working Paper 2009-01.

Williams, M. 2010. Government Cash Management: Its Interaction with Other Financial Policies. *IMF Technical Notes and Manuals (Washington: International Monetary Fund)*.

Yadav, A., Sahay, A., Yadav, M. R., Bhandari, S., Yadav, A. & Sahay, K. B. One hour Ahead Short-Term Electricity Price Forecasting Using ANN Algorithms. 2018 International Conference and Utility Exhibition on Green Energy for Sustainable Development (ICUE), 2018. IEEE, 1-4.

Yan, D., Zhou, J., Zhao, Y. & Wu, B. Short-Term Subway Passenger Flow Prediction Based on ARIMA. International Conference on Geo-Spatial Knowledge and Intelligence, 2017. Springer, 464-479.

Yan, H. & Zou, Z. 2013. Application of a Hybrid ARIMA and Neural Network Model to Water Quality Time Series Forecasting. *Journal of Convergence Information Technology,* 8.

Yan, X. & Zhao, J. Application of Neural Network in National Economic Forecast. 2018 IEEE 3rd International Conference on Image, Vision and Computing (ICIVC), 2018. IEEE, 931-935.

Yatiyana, E., Rajakaruna, S. & Ghosh, A. Wind speed and direction forecasting for wind power generation using ARIMA model. 2017 Australasian Universities Power Engineering Conference (AUPEC), 2017. IEEE, 1-6.

Yildiz, B. & Yezegel, A. 2010. Fundamental Analysis with Artificial Neural Network. *The international journal of business and finance research,* 4**,** 149.

Yip, H.-L., Fan, H. & Chiang, Y.-H. 2014. Predicting the maintenance cost of construction equipment: Comparison between general regression neural network and Box–Jenkins time series models. *Automation in Construction,* 38**,** 30-38.

Yu, L., Zhou, L., Tan, L., Jiang, H., Wang, Y., Wei, S. & Nie, S. 2014. Application of a new hybrid model with seasonal auto-regressive integrated moving average (ARIMA) and nonlinear auto-regressive neural network (NARNN) in forecasting incidence cases of HFMD in Shenzhen, China. *PloS one,* 9.

Yukseltan, E., Yucekaya, A. & Bilge, A. H. 2017. Forecasting electricity demand for Turkey: Modeling periodic variations and demand segregation. *Applied energy,* 193**,** 287-296.

Zeng, Z., Zhang, H., Zhang, R. & Yin, C. 2015. A novel feature selection method considering feature interaction. *Pattern Recognition,* 48**,** 2656-2666.

Zhang, D., Zang, G., Li, J., Ma, K. & Liu, H. 2018a. Prediction of soybean price in China using QR-RBF neural network model. *Computers and electronics in agriculture,* 154**,** 10-17.

Zhang, G., Patuwo, B. E. & Hu, M. Y. 1998. Forecasting with artificial neural networks: The state of the art. *International Journal of Forecasting,* 14**,** 35-62.

Zhang, G. P., Patuwo, B. E. & Hu, M. Y. 2001. A simulation study of artificial neural networks for nonlinear time-series forecasting. *Computers & Operations Research,* 28**,** 381-396.

Zhang, J., Wei, Y.-M., Li, D., Tan, Z. & Zhou, J. 2018b. Short term electricity load forecasting using a hybrid model. *Energy,* 158**,** 774-781.

Zhang, P. G. 2003. Time series forecasting using a hybrid ARIMA and neural network model. *Neurocomputing,* 50**,** 159-175.

Zhang, R., Xu, M., Han, M. & Li, H. Multivariate chaotic time series prediction based on improved extreme learning machine. 2017 36th Chinese Control Conference (CCC), 26-28 July 2017 2017. 4006-4011.

Zhao, Z., Chen, W., Wu, X., Chen, P. C. & Liu, J. 2017. LSTM network: a deep learning approach for short-term traffic forecast. *IET Intelligent Transport Systems,* 11**,** 68-75.

Zhou, L., Yu, L., Wang, Y., Lu, Z., Tian, L., Tan, L., Shi, Y., Nie, S. & Liu, L. 2014. A hybrid model for predicting the prevalence of schistosomiasis in humans of Qianjiang City, China. *PloS one,* 9**,** e104875.